

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 6, 2004, 18:18:04 ; Search time 833 Seconds

(without alignments)
4584.790 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRPGIWPALLGIVLAWL.....AVAGLQPDAGLFTVSEAC 899

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3373863 segs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DRV=xlh

-Q=/cgn2_1/USFTO.spool/US09970944/runat_05102004_112009_1616/app.query.fasta_1.1095

-DB=N Geneseq 29Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi

-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTPMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09970944.@CGN_1_1_581.@runat_05102004_112009_1616 -NCPU=6 -ICPU=3

-NO MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 29Jan04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002s.*

7: Geneseqn2003as.*

8: Geneseqn2003bs.*

9: Geneseqn2003cs.*

10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	4787	100.0	2881	6	ABK49422 DNA encod
2	4698.5	98.2	2752	6	ABK37922 cDNA enco
3	4688.5	97.9	2697	6	ABK52891 Human net
4	4545.5	95.0	2697	6	AAS16843 Rat retri
5	4545.5	95.0	3014	2	AAV52940 Rat UNC-5
6	4434	92.6	2907	4	AAK52261 Human pol
7	4379.5	91.5	3580	6	ABK15169 Human REP
8	2694	56.3	1787	2	AAV52941 Human UNC

9	2673.5	55.8	3646	5	AAS75738	DNA encod
10	2490.5	52.0	2955	6	ABK32062	DNA encod
11	2484	51.9	2860	6	ABT06279	Human NOV
12	2484	51.9	2860	6	ABT06280	Human NOV
13	2479	51.8	3884	4	AAS21316	Human cDN
14	2479	51.8	3884	7	ACD23925	Novel hum
15	2479	51.8	3884	7	ACA67066	cDNA enco
16	2479	51.8	3884	7	ACA03675	cDNA enco
17	2479	51.8	3884	7	ABX89213	DNA encod
18	2479	51.8	3884	7	ACD41867	Human sec
19	2479	51.8	3884	7	ACA04096	Human cDN
20	2479	51.8	3884	8	ADA45664	Novel hum
21	2479	51.8	3884	8	ADA76095	Human PRO
22	2479	51.8	3884	8	ADA18745	Human PRO
23	2479	51.8	3884	8	ADA61368	Homo sapi
24	2479	51.8	3884	8	ADB19153	Novel hum
25	2479	51.8	3884	8	ADB27694	cDNA enco
26	2479	51.8	3884	8	ADA86173	Novel hum
27	2479	51.8	3884	8	ADB15737	Human PRO
28	2479	51.8	3884	8	ADA47523	Human PRO
29	2479	51.8	3884	8	ADA67318	Human PRO
30	2479	51.8	3884	8	ADB30325	cDNA enco
31	2479	51.8	3884	8	ADA85621	Novel hum
32	2479	51.8	3884	8	ADA96833	Human PRO
33	2479	51.8	3884	8	ADA79137	Human PRO
34	2479	51.8	3884	8	ADA87276	Novel hum
35	2479	51.8	3884	8	ADB16478	Human PRO
36	2479	51.8	3884	8	ADA91570	Novel hum
37	2479	51.8	3884	8	ADB14633	Human PRO
38	2479	51.8	3884	8	ADB18594	Novel hum
39	2479	51.8	3884	8	ADA93809	Human PRO
40	2479	51.8	3884	8	ADB19705	Novel hum
41	2479	51.8	3884	8	ADB13017	Human PRO
42	2479	51.8	3884	8	ACD98496	Novel hum
43	2479	51.8	3884	8	ADA74271	Human PRO
44	2479	51.8	3884	8	ADB24504	Human PRO
45	2479	51.8	3884	8	ADA82028	Human PRO

ALIGNMENTS

RESULT 1

ABK49422

ID ABK49422 standard; DNA; 2881 BP.

XX

AC ABK49422;

XX

DT 15-JUL-2002 (first entry)

XX

DE DNA encoding human UNC5-like protein NOV1.

XX

KW Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;

KW cell signal processing; metabolic pathway modulation; cancerous tissue;

KW antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;

KW chromosome 13; gene; ds.

XX Homo sapiens.

OS

XX

FH Key

FT CDS

FT

FT

FT

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Location/Qualifiers

87..2786

/*tag= a

/product= "Human UNC5-like protein NOV1"

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

```

PI Herrmann JL, Rastelli L, Shimkets RA;
XX WPI; 2002-340104/37.
DR DR P-PSDB; AAU79939.
XX
PT Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for
PT treating cardiomyopathy, atherosclerosis, and cancer.
XX
PS Claim 8; Page 7-8; 180pp; English.
XX
CC The present invention relates to a new NOVX polypeptide having a 900
CC (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or 331 (NOV6)
CC residue amino acid sequence, as given in the specification. The novel
CC polypeptide, and its encoding polynucleotide, are used to treat
CC cardiomyopathy, atherosclerosis, cancer or a disease related to cell
CC signal processing and metabolic pathway modulation, in a human. Detecting
CC the polypeptide or polynucleotide is useful for identifying cancerous
CC tissue. The antibody can be used to treat diabetes or cancer. The host
CC cells can be used to produce non-human transgenic animals useful in drug
CC screening. The present nucleic acid sequence is that of the human UNC5-
CC like NOV1 gene located on chromosome 13. This sequence encodes the human
XX UNC5-like protein NOV1 of the invention
XX
SQ Sequence 2881 BP; 526 A; 985 C; 868 G; 502 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 2881
Score: 4787.00 Matches: 899
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-970-944-2 (1-899) x ABK49422 (1-2881)
QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAaTrpLeu 20
DB 87 ATGGCCGTCGGCCCGCGCTGTGGCCAGCGCTCTGGCCATAGTCTTGGCGCTGGCTC 146
QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
DB 147 CGCGGCTCGGGTCCAGCAGAGTGCACCGTGGCCCAACCCAGTGCCTGGTGCACCCG 205
QY 41 AspleuLeuProHisPheLeuValGluProGluAspValTrileValLysAsnLysPro 60
DB 207 GACCTGCTTCCCACTTCCCTGGTGGAGCCCGAGGATGTACATCGTCAAGAACAGCCA 265
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
DB 267 GTGCTGCTTGTGCAAGCGGTGCCCGCCAGCAGATCTTCAAGTGCACCGGGAG 326
QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
DB 327 TGGGTGCGCCAGGTGGACACAGTGTATCGAGCGCAGCAGACAGCGGAGCAGTGTGG 386
QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
DB 387 ACCATGAGGTGCCATTATGTCCTCAAGGCGAGGTCGAGAGGTTTCGGGCTGAG 446
QY 121 GluTrpTrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAla 140
DB 447 GAATACTGGTGCCAGTGGTGGATGGAGTCTCTCGGGCACCACCAAGAGTCAGAAGGCC 506
QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
DB 507 TACATCGCATAGCAGATTGCGAAGAACTTCGAGCAGGAGCGCGCTGGCCAAAGAGGTG 566
QY 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
DB 567 TCCCTGGAGCGGAGATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 626
QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTrile 200

```

```

DB 627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCGCTCCCTGACCCCAATGTATACATC 686
QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
DB 687 ACCCGGAGCAGACGCTGGTGGTGGACAGCGCGCTTCTGACACGGCAACATACACC 746
QY 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240
DB 747 TGGTGGCCAGAACATCGTGGCAGCTGCGCGAGCGCTCCGCTGCTGTCATCGTCTAC 806
QY 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
DB 807 GTGAACGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 866
QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
DB 867 GGTGGCAGAAACGAGCGCGAGCTGCACCAACCGCGCTCTCAACGGGGGCGCTTTC 926
QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
DB 927 TGTGAGGGGAGAAATGTCATGACCGACCGCTCTCTCTCTGCTGCTGCTGCTGCTGCTG 986
QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
DB 987 AGCTGGAGCCGCTGGAGCAAGTGGTGGCTGGTGGCTGGTGGCTGGTGGCTGGTGGCTGG 1046
QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGlnGlyThrAspLeu 340
DB 1047 CGTGAGTGTCTTACCACGACCCCGCAGCGAGGGGAGGAGTGCAGGCGACTGACCTG 1106
QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
DB 1107 GACACCGCAACTGTACAGTGCCTGTGTACAGTGTCTTGGCCCTGAGGAGCTG 1166
QY 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
DB 1167 GCCTCTATGTGGCGCTCATCGCGCTGGCGCTGCTGGCTGGCTGGCTGGCTGGCTGGCT 1226
QY 381 IleLeuValTyrCysArgLysGlyGlyLeuAspSerAspValAlaAspSerSerIle 400
DB 1227 ATCCTCGTATTATTCGCGAAGAGGAGGCTGAGCTCAGATGGTGGTGGTGGTGGTGGT 1286
QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
DB 1287 CTACCTCAGCTTCCAGCCCGTCCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTC 1346
QY 421 LeuThrIleGlnProAspLeuSerThrThrThrThrThrThrThrThrThrThrThrThr 440
DB 1347 CTCACCATCCAGCGGACCTCAGCACCCACCCACCTTACCAGGGCAGTCTGTGCTCCCGG 1406
QY 441 GlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGly 460
DB 1407 CAGATGGGCGCCAGCCCAAGTTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 1466
QY 461 GlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSer 480
DB 1467 GCGCGCGCCACACACTGTCACACAGCTCTCCACCTCTCAGGCGCGAGGAGTTCGTCTCC 1526
QY 481 ArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyr 500
DB 1527 CGCTCTCCACCCAGAACTACTTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 1586
QY 501 GlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeu 520
DB 1587 GGGACCTTCAACTTCTTCGGGGCGCGCTGATGATCCCTAATACAGGTATCAGGCTCTCTC 1646
QY 521 IleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLys 540
DB 1647 ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAGCTGCACAA 1706
QY 541 ProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSer 560
DB 1707 CCGAAGACGTGAGTTGCCCTAGCTAGTGGTGTGAGACCCCTGGTGTGAGTCCATCGTTAGC 1766

```


QY	561	CysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGly	580
Db	1767	TGTGGACCCCTGGCGTCTCTGCTCACCGGCCAGTCATCTTGGCTATGGACCACTGTGGG	1826
QY	581	GluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrp	600
Db	1827	GAGCCACGCCCTGACAGCTGGAGCCTCGCCTCAAAAAGCAGTCGTGCGAGGCGAGCTGG	1886
QY	601	GluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeu	620
Db	1887	GAGCAGGATGTCTGCACTCTGGGCGAGAGCGCCCTCCACCTCTACTACTGCCACGTG	1946
QY	621	GluAlaSerAlaCysTrpValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGlu	640
Db	1947	GAGGCCAGTGCCTGCTACGTCCTCACCGAGCAGCTGGCGCGCTTGGCTGGTGGAGAG	2006
QY	641	AlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCys	660
Db	2007	GCCTCAGCTGGCTGCGCGCAAGCGCCTCAAGCTGCTTCTGTTGGCGCGGTGGCGTCG	2066
QY	661	ThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLys	680
Db	2067	ACCTCCCTCGAGTACACATCCGGTCTACTGCTCGCATGACCCACGATGCACTCAAG	2126
QY	681	GluValValGlnLeuGluLysGlnLeuLysGlyGlnLeuLeuGlnGluProArgValLeu	700
Db	2127	GAGTGTGTGAGCTGGAGAACGACGTGGGGGACAGCTGATCCAGGAGCCACGGGTCCGT	2186
QY	701	HisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeu	720
Db	2187	CACCTTCAAGACAGTTACCAACCTCGCGCTATCCATCCAGATGCGCCAGCTCCCTG	2246
QY	721	TrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGly	740
Db	2247	TGGAAGAGTAAGCTCCTGTGTGAGCTACAGAGATCCCTTTTATCATCTTGAATGCG	2306
QY	741	ThrGluArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAsp	760
Db	2307	ACCGACGGGTACTTGCACTGCACCTTCACCCTGGAGGCTGTACGCCGCCACACTAGTGAC	2366
QY	761	LeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsn	780
Db	2367	CTGGCTCTGAAGCTGTGGGTGTGGCAGTGTGAGGCGACGGCAGAGCTTCAGCATCAAC	2426
QY	781	PheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly	800
Db	2427	TTCAACATCAACCAAGGACACAAGGTTTGTCTGAGCTGTCTGCTCTGGAGAGTGAAGCGGG	2486
QY	801	ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIle	820
Db	2487	GTCCCGACCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAGATTA	2546
QY	821	IleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLys	840
Db	2547	ATTTCAGCGCTGACCCACCCCTGTAGCGGGGTGCGACCTGGCGGAGCTCTGGGCCAGAAA	2606
QY	841	LeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIle	860
Db	2607	CTCCACTGGACAGCCATCTACAGCTCTTTGGCTTCCAGGCCAGCCCCCAGGCCATGATC	2666
QY	861	LeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla	880
Db	2667	CTCAACTGTGGGAGCGCGGCATCTCCCAACGGCAACCTCAGCCAGCTGGCTGGCAGCA	2726
QY	881	ValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys	899
Db	2727	GTGGCTGGACTGGGCCACGACAGCTGGCGCTCTTACAGTGTGGAGGCTGAGTGC	2783

RESULT 2

ABK37922

ID ABK37922 standard; cDNA; 2752 BP.

XX

ABK37922;

21-MAY-2002 (first entry)

cdDNA encoding Human protein NOV1.

Human: NOVX; ss; gene; cardiomyopathy; atherosclerosis; diabetes; cell signal processing disorder; metabolic disorder; obesity; infection; anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder; Alzheimer's disease; Parkinson's disease; immune disorder; haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension; osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris; myocardial infarction; ulcer; allergy; benign prostatic hypertrophy; psychosis; neurological disorder; anxiety; schizophrenia; manic depression; dementia; dyskinesia; Huntington's disease; Gilles de la Tourette's syndrome; gene therapy.

Homo sapiens.

WO200210216-A2.

07-FEB-2002

30-JUL-2001: 2001WO-IIS024225

28-JUL-2000: 2000US-0221409D

28-JUL-2000; 2000US-0221409E.
04-AUG-2000: 2000US-0222840P

04-AUG-2000; 2000US-0222640P.
04-AUG-2000: 2000US-0223752P.

04-AUG-2000: 2000US-0223762P.
04-AUG-2000: 2000US-0223732E.

01 AUG 2000; 2000US-0223702E;
04-AUG-2000; 2000US-0223769P;

04-AUG-2000; 2000US-0223770P.

14-AUG-2000; 2000US-0225146P.

15-AUG-2000; 2000US-0225392P.

15-AUG-2000; 2000US-0225470P.

16-AUG-2000; 2000US-0225697P.

01-FEB-2001; 2001US-0263662P.

05-APR-2001; 2001US-0281645P.

(CURA-) CURAGEN CORP.

Padigaru M, Mezes P, Mishra V, Burgess C, Casman S, Grosse WM, Alsobrook JP, Lepley DM, Gerlach VL, Macdougall JR, Smithson G, G; G

WPI; 2002-180074/23.

P-PSDB; AAU85403.

New isolated cytoplasmic, nuclear, membrane bound, or secreted polypeptide, useful for treating cardiomyopathy, atherosclerosis, infections, cancer, neurodegenerative, metabolic, hematopoietic and immune disorders.

Claim 9: Page 9-10; 213pp; English.

The invention relates to an isolated cytoplasmic, nuclear, membrane bound, or secreted polypeptide (NOVX, x= 1-14) their variants or mature form. Also included are the nucleic acids encoding the NOVX proteins, a vector comprising the nucleic acid, a cell comprising the vector, an anti-NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the antibody are useful for treating or preventing a NOVX-associated disorder, where the disorder is selected from cardiomyopathy, atherosclerosis, diabetes, a disorder related to cell signal processing and metabolic pathway modulation, metabolic disorders, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, immune disorders, haematopoietic disorders, and the various dyslipidaemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases, bacterial, fungal, protozoal and viral infections, pain, bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's disease, multiple sclerosis, Albright Hereditary Osteodystrophy, angina pectoris, myocardial infarction, ulcer, allergy, benign prostatic hyper trophy, and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, and dyskinesias.

CC such as Huntington's disease and Gilles de la Tourette's syndrome. The
 CC nucleic acid is useful in gene therapy. The present sequence encodes a
 CC NOVX protein

XX
 SQ Sequence 2752 BP; 505 A; 937 C; 829 G; 481 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 9,21e-315 Length: 2752
 Score: 4698.50 Matches: 888
 Percent Similarity: 98.89% Conservative: 2
 Best Local Similarity: 98.67% Mismatches: 7
 Query Match: 98.15% Indels: 3
 DB: 6 Gaps: 3

US-09-970-944-2 (1-899) x ABK37922 (1-2752)

Qy 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAaTrpLeu 20
 Db 46 ATGGCGGTCCGGCCCGCCCTGTGGCCAGCGCTCTGGGCATAGTCTCGCGCTTGCTC 105
 Qy 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
 Db 106 CGGGCTCGGGTGGCCAGCAGAGTGCCACCGTGGCCCAACCCAGTGCCTGTGGCAACCCG 165
 Qy 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
 Db 166 GACCTGCTTCCCACTTCTCGTGGAGCCCGAGGATGTGTACATCTGTCAAGTGCAACGGGAG 225
 Qy 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
 Db 226 GTGCTGCTGTGTGCAAGGCGGTGCCCGCCAGCAGATCTTCAAGTGCAACGGGAG 285
 Qy 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
 Db 286 TGGGTGGCCAGGTGGACCACTGATCGAGCGCAGCACAGCGGAGCAGTGGGCTGCC 345
 Qy 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
 Db 346 ACCATGAGGTGCCGATTAATGTCTCAAGGAGCAGGTGCAAGAGGTGTTCCGGCTGGAG 405
 Qy 121 GluTyrTrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAla 140
 Db 406 GAATACTGGTGCCAGTGGCTGGCATGGAGCTCTCGGGCACCAAGAGTCAAGAGGCC 465
 Qy 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
 Db 466 TACATCGGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCGCGCTGGCCAAAGAGGTG 525
 Qy 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
 Db 526 TCCCTGGAGCAGGGCATCGTGTGCTGCCCTGGCTCCACGGAGGGCATCCCTCCAGCCGAG 585
 Qy 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
 Db 586 GTGGAGTGGCTCCGGAACAGGAGCCTGTGTGACCCGCTCCCTGGACCCCAATGTATACATC 645
 Qy 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
 Db 646 ACGGGGAGCAGACCCCTGGTGGTGGCAGACCGCCGCTGTGTGACCGGCCCACTACACC 705
 Qy 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240
 Db 706 TGGTGGCCCAAGAACATGTGGCAGCTGCCCGCAGCGGCTCCGCTGTGTATCGTCTAC 765
 Qy 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
 Db 766 GTGAACGGTGGTGGTTCGACCTGGACCGAGTGGTCCGCTCTGAGCGGCGCAGCTGTGGCGC 825
 Qy 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyValPhe 280
 Db 826 GGCTGGCAGAACCGGAGCGGAGTGCACCAACCGCGGCTCTCAACGGGGGCGCTTTC 885
 Qy 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300

Db 886 TGTGAGGGGAGATGTC---CAGAAAACAGCCTGGCCACCTGTGCCAGTAGACGGC 942
 Qy 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
 Db 943 AGCTGGAGCCCGTGGAGCAAGTGTGGCCCTGTGGCTGGACTGCACCCAGCTGGCGGAGC 1002
 Qy 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGlnGlyThrAspLeu 340
 Db 1003 CGTGAGTGTCTTACCAGCACCCCGCAACGGAGGGAGAGTGCAGGAGCTAGACCTG 1062
 Qy 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
 Db 1063 GACACCCGCAACTGTACCACTGCTGTGTACACAGTCTCTGTGSCCTGAGGAGCTG 1122
 Qy 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
 Db 1123 GCCCTCTATGTGGCCCTCATCGCGTGGCGCTTGCCTGTGCTGCTGCTGCTGCTC 1182
 Qy 381 IleLeuValTyrCysArgLysLysGluCysLeuAspSerAspValAlaAspSerSerIle 400
 Db 1183 ATCTCTGTTTATTGCCGGAAGAGGAGGGCTGGACTCAGATGGGCTGACTGCTGCTCAT 1242
 Qy 401 LeuThrSerGlyPheGlnProValSerIleIlePheProSerLysAlaAspAsnProHisLeu 420
 Db 1243 CTACCTCAGGCTTCCAGCCCGTCCAGCATCAAGCCAGCAAGAGCAGCAACCCCATCTG 1302
 Qy 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrThrGlnGlySerLeuCysPro 439
 Db 1303 CTCACATCCAGCGGACCTCAGCACCCACCACCACCTACCCAGGGCAGTCTCTGCTCC 1362
 Qy 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
 Db 1363 CGGAGGATGGGCGCCAGCCCAAGTTCAGCTCACCATGGGCACCTGTCTCAGCCCCCTG 1422
 Qy 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
 Db 1423 GGTGGCGCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCGGAGGAGTCTGCTC 1482
 Qy 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
 Db 1483 TCCGCGCTCTCCACCCAGAACTACTCTCGCTCCCTGCGCCGAGGACACAGCAATGACC 1542
 Qy 500 TyrGlyThrPheAsnPheLeuGlyValArgLeuMetIleProAsnThrGlyIleSerLeu 519
 Db 1543 TATGGACCTTCAACTTCTCGGGGCGCGCTGATGATCCCTAATACAGGTATCAGCCTC 1602
 Qy 520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
 Db 1603 CTATATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCAGCTGCAC 1662
 Qy 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
 Db 1663 AAGCGGAAGACGTGAGGTGGCCCTAGCTGGCTGTGAGACCCCTGCTGAGTCCATCGTT 1722
 Qy 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
 Db 1723 AGCTGTGACCCCTGGCTGCTGCTCATCCCGGCGAGTCTCTCTGGCTATGGACCATGT 1782
 Qy 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
 Db 1783 GGGGAGCCAGCCCTGACAGCTGGAGCTGGCGCTCAAAAAGACAGTCTGTGAGGGGAGC 1842
 Qy 600 TrpGlnGlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGln 619
 Db 1843 TGGGAG---GATGTGCTGCACCTGGGAGGAGGCGGCTCCACCTCTACTACTGCCAG 1899
 Qy 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
 Db 1900 CTGGAGGCGAGTGTCTGTCTGTCTTTCACCGAGCAGCTGGGCGCTTTCCTGCTGGTGA 1959
 Qy 640 GluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659

```
Db 1960 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTGTTTGGCGCGGTGCC 2019
Qy CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db TGCACCTCCCTCGAGTCAACATCGGGTCTACTGCTGTCATGACCCACGATGCACTC 2079
Qy 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuGlnGluProArgVal 699
Db 2080 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTC 2139
Qy 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db 2140 CTGCACCTCAAGGACAGTTACCAACACCTGCGCTATCCATCCAGATGTGCCAGCTCC 2199
Qy 720 LeuTyrLysSerLysLeuLeuValSerTyrGlnGluLeuProPheTyrHisIleTyrAsn 739
Db 2200 CTGTGGAGAGTAGCTCTTGTGAGTACGAGAGATCCCTTTTATCATCTTGAAT 2259
Qy 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db 2260 GGCACGCGAGCGGTACTTGCACTGCACCTTCACCTGGAGCGGTGCAGCCCGACCTAGT 2319
Qy 760 AspLeuAlaCysLysLeuTyrValTyrGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db 2320 GACCTGGCCTGCAAGCTGTGGTGTGGCAGTGGAGGGCGACGGGCGAGAGCTTCAGCATC 2379
Qy 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuAlaLeuGluSerGluAla 799
Db 2380 AACTTCACATCACAGGACACNAGGTTTGTGAGCTGTGGCTCTGGAGAGTGAAGCG 2439
Qy 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db 2440 GGGGTCCCGAGCCCTGGTGGGCCCCAGNGCTTCAAGATCCCTTCCTCATTCGGCAGAAG 2499
Qy 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTyrArgThrLeuAlaGln 839
Db 2500 ATAATTTCCAGCTGGAGCCACCCCTGTGAGCGGGGTGCCAGTGGCGGACTCTGGCCCGAG 2559
Qy 840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db 2560 AACTCCACTGGACAGCCATCTCAGCTTCTTTCCTCCAAAGCCCGCCAGCCCATG 2619
Qy 860 IleLeuAsnLeuTyrGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db 2620 ATCTCAACTGTGGGAGGCGGCGCACTTCCCAACGGCACTTCAGCCAGCTGGCTGCA 2679
Qy 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 2680 GCAGTGGCTGGAGTGGGCCAGCCAGCGCTGGCTCTTTCACAGTGTGGAGGCTGAGTGC 2739

RESULT 3
ABK52891
ID ABK52891 standard; DNA; 2697 BP.
XX
AC ABK52891;
XX
DT 27-AUG-2002 (first entry)
DE
DE Human netrin binding membrane receptor UNC5H-1 DNA sequence #1.
XX
KW Netrin binding membrane receptor; receptor; UNC5H-1; gene; ds; human;
KW neurotropic; neuroprotective; cytoskeletal; antiparkinsonian;
KW cerebroprotective; cancer; central nervous system; CNS; stroke;
KW Parkinson's disease; multiple sclerosis; Alzheimer's disease.
XX
OS Homo sapiens.
XX
FH Key
FH CDS
FT 1..2697
FT /*tag= a
FT /product= "Netrin binding membrane receptor UNC5H-1"
XX
XX W0200233080-A2.
```

```
XX 25-APR-2002.
PD
XX 15-OCT-2001; 2001WO-EP011891.
PF
XX 16-OCT-2000; 2000US-0240061P.
PR
XX (FARB ) BAYER AG.
PA
XX Koehler RH;
PI
XX WPI; 2002-463314/49.
DR
XX P-PSDB; AAU97899.
DR
XX Novel human netrin binding membrane receptor polypeptide and
PT polynucleotides for identifying modulating agents useful in treating
PT diseases e.g. Parkinson's disease, multiple sclerosis, stroke,
PT Alzheimer's disease.
XX
XX Claim 1; Fig 1; 94pp; English.
XX
CC This invention relates to the DNA and protein sequences of a novel
CC purified human netrin binding membrane receptor, UNC5H-1. The DNA
CC sequence of the invention is useful as a probe for detecting a nucleic
CC acid encoding the UNC5H-1 protein in a biological sample. The sequences
CC of the invention are useful to screen for agents which decrease the
CC activity of the UNC5H-1 protein. The sequences are also useful for
CC screening agents which regulate (modulate) the activity of the protein of
CC the invention. A pharmaceutical composition containing the protein of the
CC invention or a reagent that modulates the activity of the UNC5H-1 protein
CC may be useful for treating a UNC5H-1 dysfunction related disease such as
CC cancer or a central nervous system (CNS) disorders (e.g. Parkinson's
CC disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion
CC proteins comprising the UNC5H-1 protein are useful for generating
CC antibodies and for in various assay systems, and the protein can be used
CC as a bait protein in a two-hybrid assay or three-hybrid assay. The method
CC of the invention is useful for detecting a coding sequence for the UNC5H-
CC 1 protein. The present sequence represents a DNA sequence encoding the
CC human netrin binding membrane receptor UNC5H-1 protein of the invention
XX
XX Sequence 2697 BP; 503 A; 906 C; 807 G; 481 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4,4e-314 Length: 2697
Score: 4688.50 Matches: 886
Percent Similarity: 98.78% Conservative: 3
Best Local Similarity: 98.44% Mismatches: 8
Query Match: 97.94% Indels: 3
DB: Gaps: 3

US-09-970-944-2 (1-899) x ABK52891 (1-2697)
Qy 1 MetAlaValArgProGlyLeuTyrProAlaLeuLeuGlyIleValLeuAlaAAlaTyrLeu 20
Db 1 ATGGCGGTCCGGCCCGGCTGTGGCCAGCGCTCTGGGCATAGTCTCGCCGCTTGCTC 60
Qy 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db 61 CGCGGCTCGGGTGGCCAGCAGAGTGCACCGTGGCCAAACCCAGTGCCTGGTGCACACCG 120
Qy 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db 121 GACCTCTCTCCACATCTCTGTGGAGCCCGAGGATGTGTACATCTCTCAAGAACAGCA 180
Qy 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db 181 GTGCTGCTTGTGCAAGCGCTGCCCGCCAGCAGATCTTCTCAAGTGCACCGGGAG 240
Qy 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerGlyGluPro 100
Db 241 TGGGTGCGCGAGTGGACCACTGATCGAGCGCAGCAGACGCGGAGCAGTGGCTGCC 300
Qy 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
```

Db 301 ACCATGAGGTCCGCTAATTAAGTCTCAAGGACGAGAGTTCAGAAAGGTTTCGGGCTGGAG 360
Qy 121 GluTyrTrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAla 140
Db 361 GAATACTGGTGCCAGTCCGTGGCATGAGCTCTCCGGGCACCAAGAGTCAGAAAGGCC 420
Qy 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnProLeuAlaLysGluVal 160
Db 421 TACATCCGCATAGCTATTTCGGCAAGAACTTCAGCAGAGCGCTGGCCAAAGAGGTG 480
Qy 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db 481 TCCCTGGAGCAGGCGATCGTGTCTGCTGCGTCCACCGAGGCGCATCCCTCCAGCCGAG 540
Qy 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db 541 GTGGAGTGGTCCGGAACGAGGACCTGGTGACCCGCTCCCTGGAGCCCAATGTATACATC 600
Qy 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
Db 601 ACOCGGAGCAGACCCCTGGTGGTGGCAGAGCCCGCTTGTGTGACAGCGCCAACTACACC 660
Qy 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240
Db 661 TGGTGGCCCAAGAACTCGTGGCAGCTCGCGCAGCGCTCCGTGTGTGTCATCGTCTAC 720
Qy 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
Db 721 GTAAACGGTGGTGGTGCAGCGTGACCGAGTGGTCCGTCTGTCAGCGCGCAGCTGTGGGCGC 780
Qy 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyValPhe 280
Db 781 GGCTGGCAGAAACGGACCGGAGCTGCACCAACCGCGCGCTCTCAACGGGGGCGCTTTC 840
Qy 281 CysGluGlyGlnAsnValHisAspArgValSerSerLeuValSerValAspGly 300
Db 841 TGTGAGGGGAGATGTC---CAGAAACACGCTGGCGCCACCTGTGCCAGTGGACGGC 897
Qy 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db 898 AGCTGGAGCCGTGGACAGTGGTGGCCCTGTGGGCTGGACTGCACCACTGGCGGAGC 957
Qy 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db 958 CGTGAGTGCTCTGACCCAGACCCCGCAACGGAGGGAGGAGTGCAGGGCAGCTGACCTG 1017
Qy 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db 1018 GACACCCGCAACTGTACAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTG 1077
Qy 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
Db 1078 GCCCTCTATGGGCGCTCATCGCGTGGCCGCTCTGCCTGGTCTGCTGCTGCTGCTGCTC 1137
Qy 381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
Db 1138 ATCTCTGTTATTGCCGGAAGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCAT 1197
Qy 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db 1198 CTCACCTCAGGCTTCAGCCCGCTCAGCATCAAGCCAGCAAGCAGACCAACCCCATCTG 1257
Qy 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrGlnGlnSerLeuCysPro 439
Db 1258 CTCACCATCCAGCGGACCTCAGACCAACCAACCAACCTACCGAGGGCAGTCTCTGTCCC 1317
Qy 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db 1318 CGCGAGGATGGGCCCGACCCCAAGTTCAGTCCAGTCCACCAATGGGCACCTGCTCAGCCCTG 1377
Qy 460 GlyGlyGlyArgHisThrLeuHisHisSerProThrSerGluAlaGluPheVal 479

Db 1378 GGTGGGGCGCCGACACACTGACCAACAGCTCTCCACCTCTGAGGGCCGAGGAGTTTCGTC 1437
Qy 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db 1438 TCCCGCTCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGACACCAACATGACC 1497
Qy 500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db 1498 TATGGGACCTTCAACTTCTCGGGGCGCGCTGATGATCCTTAATACAGGAATCAGCCTC 1557
Qy 520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
Db 1558 CTCACTCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCAGCTGCAC 1617
Qy 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db 1618 AAGCCGGAACGCTGAGGTTGCCCTAGCTGGCTGTGAGACCCCTGCTGAGTCCCATCGTT 1677
Qy 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db 1678 AGCTGTGACCCCTGGCGTCTGCTCACCCGGCCAGTCTATCTGGCTATGGACCACTGT 1737
Qy 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db 1738 GGGGAGCCGACGCTGACAGCTGGAGCTGGCGCTCAAAAAGCAGTCTGTCGAGGGCAGC 1797
Qy 600 TrpGluAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGln 619
Db 1798 TGGGAG---GATGTGTGCACCTGGCGAGGAGCGCCCTCCACCTTACTACTGCGAG 1854
Qy 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db 1855 CTGGAGCCAGTGGCTGCTGCTGCTTCAACCGAGCAGCTGGGCGCTTTCGCTTGGTGGGA 1914
Qy 640 GluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
Db 1915 GAGGCCCTCAGCGTGGCTGGCGCCGACAGCGCTCAAGCTGCTTCTGTTGGCCCGCTGGCC 1974
Qy 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db 1975 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCTGTGATGACACCCAGATGCATC 2034
Qy 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
Db 2035 AAGGAGTGGTGGCAGCTGGAGAACAGCTGGGGGAGCAGCTGATCCAGGACCCAGCGTCC 2094
Qy 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db 2095 CTGCACCTCAAGGACAGTTACCAACAGCTGGCGCTATCCATCCAGATGTGCCAGGTCC 2154
Qy 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db 2155 CTGTGGAGAGTAAAGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2214
Qy 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db 2215 GGCACGACGCGTACTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2274
Qy 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db 2275 GACCTGGCGCTCAAGCTGTGGGTGTGGCAGTGGAGGGCGACGGGCGAGCTTCAGCATC 2334
Qy 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db 2335 AACTTCAACATCACCAAGGACCAAGGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2394
Qy 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db 2395 GGGGTCCAGCCCTGGTGGGCCCCCAGTTCCTTCAAGATCCCTTCTCATTCGCGAGAAG 2454
Qy 820 IleIleSerSerLeuAspProProCysArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db 2455 ATAATTTCCAGCTGGACCCACCCCTGTAGGCGGGTGCCTGCTGCTGCTGCTGCTGCTGCTG 2514

QY 840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
 Db 2515 AAACCTCAGCTGGACAGCATCTCAGCTTCTTGGCTTCAAAGCCAGCCCGCCAGCCCATG 2574
 QY 860 IleLeuAsnLeuTrpGluAlaAraHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
 Db 2575 ATCTCTCAACTGTGGAGGCGCGCACTTCCCAAGCGCAACCTCAGCAGCTGGCTGCA 2634
 QY 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
 Db 2635 GCAGTGGCTGACTGGGCGCCAGCAGCTGGCTCTTCAAGTGTGGAGGCTGAGTGC 2694

RESULT 4
 AAS16843
 ID AAS16843 standard; cDNA; 2697 BP.
 XX
 AC AAS16843;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Rat netrin receptor UNC5H1 (YSG7) cDNA.

KW YSG; YSG7; schizophrenia; chronic animal model; LCGU; netrin receptor;
 KW local cerebral glucose utilisation; phosphodiesterase 1-alpha; UNC5H1;
 KW calcium-independent alpha-latrotoxin receptor; CIRL; trkE; synapsin 1A;
 KW epithelial discoidin domain receptor 1; synapsin 1B; neuroleptic; ss;
 KW tumour necrosis factor alpha; TNF-alpha; rat.

OS Rattus sp.

XX Key Location/Qualifiers
 PH 1. 2697
 FT CDS
 FT /*tag= a
 FT /product= "Rat netrin receptor UNC5H1"

FN WO200175440-A2.

XX 11-OCT-2001..

XX 02-APR-2001; 2001WO-GH001486.

XX 31-MAR-2000; 2000GB-00007880.

XX 26-MAY-2000; 2000GB-00012768.

XX (WELF-) WELFIDE CORP.

XX Cochran S, Paterson G, Ohashi Y, Morris B, Pratt J;

XX WPI; 2002-010813/01.

XX P-PSDB; AAU10543.

XX Novel chronic animal model of schizophrenia, useful for identifying anti-

XX psychotic drugs and genes that are associated with schizophrenia.

XX Claim 1; Fig 8a; 79pp; English.

XX The invention relates to YSG polynucleotide fragments for use in
 CC diagnosing and/or developing treatments for schizophrenia using chronic
 CC animal models. The polynucleotides and their encoded polypeptides are
 CC used for identification of compounds which modulate the expression of YSG
 CC molecules, leading to the manufacture of schizophrenia medicaments. The
 CC sequences can also be used for testing candidate compounds for any effect
 CC on the polypeptides. Anti-schizophrenic effects of a compound can be
 CC determined by measuring local cerebral glucose utilisation (LCGU) or
 CC comparing its expression level with that of a control group. The
 CC sequences are useful in the identification of genes associated with
 CC schizophrenia states and in the development of an antibody. The sequences
 CC of the invention include phosphodiesterase 1-alpha, calcium-independent
 CC alpha-latrotoxin receptors (CIRL)-1,2k3, epithelial discoidin domain
 CC receptor 1 (trkE), netrin receptor (UNC5H1), synapsins 1A and AB and
 CC tumour necrosis factor (TNF) alpha. This sequence represents rat netrin
 CC receptor UNC5H1 (YSG7) DNA

XX SQ Sequence 2697 BP; 541 A; 864 C; 766 G; 526 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3.32e-304 Length: 2697
 Score: 4545.50 Matches: 852
 Percent Similarity: 96.78% Conservative: 19
 Best Local Similarity: 94.67% Mismatches: 26
 Query Match: 94.96% Indels: 3
 DB: 6 Gaps: 3

US-09-970-944-2 (1-899) x AAS16843 (1-2697)

QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyLeValLeuAlaAlaTrpLeu 20
 Db 1 ATGGCCGTCGGCCCGCCGCTGTGGCCAGTGTCTCTGGGCATAGTCTCGCGCCCTGGCTT 60
 QY 21 ArgGlySerGlyAlaGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
 Db 61 CGTGGTTCGGGTGCCAGCAGAGTGCACCGTGGCCCAATCCAGTGGCCGGTGCCCAAGCCC 120
 QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
 Db 121 GACCTGCTGCCCCACTTCTCTGTAGAGCTTGAGACGCTGTACATTGTCAAGAACACGCG 180
 QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
 Db 181 GTGTGTGTGGTGTGCAAGGCTGTCCCTGCCACCATCTTCTTCAAGTGCATGGGAA 240
 QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
 Db 241 TGGGTCCGCCAGGTGCTGATCAGTAATGAACGAGCAGCCAGCAGCAGCGGATTGCCA 300
 QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
 Db 301 ACCATGGAGTCCGTATCAACGTATCGAGCAGCAGGTAGAGAAAGTGTGGCTGGAG 360
 QY 121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
 Db 361 GAATACTGGTGGCAGTGTGTGGCATGGAGTCTCGGGTACCACCAAAAGTCAGAGGCC 420
 QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
 Db 421 TACATCGGATTGCCATTTCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAAGGAAGT 480
 QY 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
 Db 481 TCACTGGAGCAAGGCATTGTACTACCTTGTGCGCCCGCCAGAGGAATCCCGCAGCTGAG 540
 QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
 Db 541 GTGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTATCATC 600
 QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
 Db 601 ACGGGAGCAGCAGCCCTAGTGTGGTCCAGCCCGCTGGCCGACACGCGCAACTACACC 660
 QY 221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
 Db 661 TGTGTGCCCAAGAACATCGTACCCGTCGCCGAGACACCTTCGACGGGTCAATGTTTAT 720
 QY 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
 Db 721 GTAAACGTTGGTGGTTCGACGTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 780
 QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
 Db 781 GCTGTGCAAAACGAGCGGAGCTGCACCAACCGCGCACCTCTCAACGCGGGCGGCTTC 840
 QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
 Db 841 TGTGAGGGGCAAGTGC---CAGAAACAGCGCTGGCCACCTCTGTGCCCGCATGGATGGG 897

301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
898 AGCTGGAGTTCTGCGGAGTAAGTGGTCAAGCTGTGGGCTTGACTGCACCCACTGCGGAGC 957
321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGlnGlyThrAspLeu 340
958 CGGAGTGCTCTGACCCAGACACCCCGCAGTGGAGGTGAGGAGTGTGGGGTGTGACCTG 1017
341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
1018 GACACCCGCAACTGTACAGTGACCTCTGCTGCACACCGCTTCTTCCGCCGAGGAGCTG 1077
361 AlaLeuTyrValGlyLeuLeuAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
1078 GCTCTCTACATCGGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1137
381 IleLeuValTyrCysArgLysLysGlyLeuAspSerAspValAlaAspSerSerIle 400
1138 GGACTCATTTACTGTCCAGAGAGGAGGCTGGACTCCGATGTGGCGGACTGCTGCATC 1197
401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
1198 CTCACCTCGGCTTCCAGCCTGTCCAGCATCAAGCCAGCAGCAGCAGCAGCAGCAGCAG 1257
421 LeuThrIleGlnProAspLeuSer--ThrThrThrThrThrThrThrThrThrThrThr 439
1258 CTCACCATTCAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1317
440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
1318 AGCAGAGATGGACCCAGCCCAAGTTCAGCTCTCTAATGTGTCACTGCTCAGCCACTG 1377
460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
1378 GGGAGTGGCGCCATACGTTGCACACACACAGCTCAGCCACCTCTGAGGCTGAGGACTTGC 1437
480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
1438 TCCGCGCTCTCCACCCCAAACTACTTTGTTTCCCTGCGCGCGGACCCAGCAACATGGCC 1497
500 TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
1498 TAGGGGACCTTCACTTCTCGGGGGCGGCTGATGATCCCTAATACGGGATCAGCCTC 1557
520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
1558 CTCATACCCCGGATGCCATCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617
540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
1618 AAGCCAGAGACGTGAGGTTCGCCCTAGCTGGCTGTGAGACCCCTGCTGAGTCCAGTGGT 1677
560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
1678 AGCTGTGGCCCCCAGGAGTCTCTGCTACCCGCGGCGAGTCATCTTGGCAATGGACCACTGT 1737
580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
1738 GGAGAGCCACCCCTGACAGCTGGAGTCTGGCCCTCAAAAAGCAGTCTCGCGGGGAGT 1797
600 TrpGluGlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGln 619
1798 TGGAG---GAITGCTGCACCTTGGTGAGGAGTACCTTCCACCTCTACTACTGCGCAG 1854
620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
1855 CTGGAGGCGGGGCTGTATGTCTTCACGAGCAGTGGCGGCTTGGCCCTGTGTAGGA 1914
640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
1915 GAGGCGCTCAGCGTGGCTGCCACCAAGCGCTCAGGCTCTCTCTTGTCTCCCGTGGCC 1974
660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679

1775 TGTAGCTCCCTTGGTACAACTCCGAGTGTACTGCTACACGACACCCACGAGCTCTC 2034
680 LysGluValValGlnLeuLeuLysGlnLeuGlyGlnLeuLeuGlnGluProArgVal 699
2035 AAGGAGTGTGTCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGGCTC 2094
700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisaspValProSerSer 719
2095 CTGCACCTTCAAGACAGCTTACCAACACCTACGCTCTCCATCCACGAGTGGCCAGCTCC 2154
720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
2155 CTGTGGAGAGCAAGCTACTTGTGACGTACAGGAGATCCCTTTTACCACTCTGGAAC 2214
740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
2215 GGCACCCAGCAGTATCTGCACCTTCACCTTGGAGCGCATCAACGCCAGCAGCAGCAGC 2274
760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
2275 GACCTGCGCTGCAAGGTGTGGTGTGSCAGGTGGAGGAGATGGCAGAGCTTCAACATC 2334
780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
2335 AACTTCAACATCACTAAGGACACAAAGCTTGTGCTGATTTGCTGTGGAGGTGAAGG 2394
800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
2395 GGGTCCAGCAGCTGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAAAAG 2454
820 IleIleSerSerLeuAspProProCysArgArgGlyValaAspTrpArgThrLeuAlaGln 839
2455 ATCATCCCGAGTCTGGACCCACCTCTGCAGCGGGGCGCCGAGCTGAGAGACTCTAGCCAG 2514
840 LysLeuHisLeuAspSerHisLeuSerPheAlaSerLysProSerProThrAlaMet 859
2515 AACTTCACTGGACAGCCATCTTAGCTTCTTTCCTCCAAAGCCAGCCCTTACAGCCATG 2574
860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
2575 ATCTCAACCTATGGAGGCGACGCGACTTCCCCAACGCGCAACCTCGGCGCAGCTGCGACA 2634
880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
2635 GCTGTGGCGGACTGGGCGCAACCAAGATGCTGGCCCTCTTCACGGTGTGGAGGCGGAGTGT 2694
RESULT 5
AAV52940
ID AAV52940 standard; cDNA; 3014 BP.
XX
AC AAV52940;
XX
DT 25-MAR-2003 (revised)
DT 21-DEC-1998 (first entry)
XX
DE Rat UNC-5 homologue unc5h-1 cDNA.
XX
KW UNC-5; UNC5H-1; rat; netrin receptor; cell migration; axon guidance;
KW diagnosis; therapy; ds.
XX
OS Rattus sp.
XX
FH Key Location/Qualifiers
FT CDS 1..2697
FT /*tag= a
XX
PN WO9837085-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-US003143.
XX

PR 19-FEB-1997; 97US-00808982.
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasu K;
 XX
 DR WPI; 1998-495364/42.
 XX P-ESDB; AAW78898.
 XX
 PT Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and
 PT the biopharmaceutical industry.
 XX
 PS Claim 7; Page 15-17; 32pp; English.
 XX
 CC This cDNA, termed unc5h-1, comprises a rat homologue of *Caenorhabditis*
 CC *elegans* unc-5. Rat unc5h-1 and unc5h-2 (see AAV52942) cDNAs were isolated
 CC from an E18 brain cDNA library. The predicted proteins (see AAW78898 and
 CC AAW78900) show similarity with UNC-5. They are predicted to be involved
 CC in cell migration and axon guidance, and are characterised as receptor
 CC proteins for netrins. Gene expression is observed in regions where
 CC differentiating neurons are undergoing axogenesis. Human unc5h-1 (see
 CC AAV52941) and unc5h-2 (see AAV52943) cDNAs are also claimed. Vertebrate
 CC UNC-5 proteins may be produced recombinantly from transfected host cells
 CC by utilising these vertebrate UNC-5 nucleic acids. The invention also
 CC provides unc-5 hybridisation probes and primers, vertebrate UNC-5-
 CC specific binding agents such as specific antibodies, and methods of
 CC making and using the subject compositions in diagnosis (e.g. genetic
 CC hybridisation screens for vertebrate unc-5 transcripts), therapy (e.g.
 CC gene therapy to modulate vertebrate unc-5 gene expression) and in the
 CC biopharmaceutical industry (e.g. as immunogens, reagents for modulating
 CC cell guidance, reagents for screening chemical libraries for lead
 CC pharmacological agents, etc.). (Updated on 25-MAR-2003 to correct PI
 CC field.)
 XX
 SQ Sequence 3014 BP; 596 A; 977 C; 849 G; 592 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3,84e-304 Length: 3014
 Score: 4545.50 Matches: 852
 Percent Similarity: 96.78% Conservativity: 19
 Best Local Similarity: 94.67% Mismatches: 26
 Query Match: 94.96% Indels: 3
 DB: 2 Gaps: 3

US-09-970-944-2 (1-899) x AAV52940 (1-3014)

QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
 DB 1 ATGGCGGTCCGGCCCGGCGCTGTGGCCAGTGTCTCTGGGCATAGTCTCTCGCGCTGGCTT 60

QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
 DB 61 CTGTGGTTCGGGTGCCAGCAGAGTGCACCGTGGCCCAATCCAGTGGCCGGTGCCAAACCCC 120

QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValIysAsnLysPro 60
 DB 121 GACCTGTGCTGCCACATCTCTGTGTAGAGCCCTGAGGAGCTGTACATTGTCAAGACACAGCCG 180

QY 61 ValLeuLeuValCysIysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
 DB 181 GTGTGTGTGGTGGCAAGGCTGTGCTGCCACCCAGATCTTCTCAAGTGCATGGGAA 240

QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
 DB 241 TGGGTCCGCAGGTGCATACGTAAATTTGAACGACGACCCAGCACCGACGACGCGGATTCGCCA 300

QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluIysValPheGlyLeuGlu 120
 DB 301 ACCATGGAGGTGGTATCAACGTATCGAGGCACAGGTAGAGAAAGTGTGGGCTGGAG 360

QY 121 GluTyrTrpCysGlnCysValAlaIaTrpSerSerSerGlyThrThrLysSerGlnIysAla 140
 DB 361 GAATACTGTGGCAGTGTGTGGCATGGAGTCTCTCGGTACCAACCAAAAGTCAGAGGCC 420

QY 141 TyrIleArgIleAlaArgLeuArgIysAsnPheGluGlnGluProLeuAlaIysGluVal 160
 DB 421 TACATCCGGATTGGCTATTTCGCAAGAACTTTGAGCAGGAGGCCACTGCGCAAGAGTG 480

QY 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
 DB 481 TCATCTGGAGCAAGGCAATTGTACTACCTTGTGCGCCCCCAGAGGAATCCCCCACTGAG 540

QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
 DB 541 GTGAGTGGCTTCGAAATGAGGACCTCGTGACCCCTCCCTCGATCCCAATGTGTACATC 600

QY 201 ThrArgGlnHisSerLeuValValArgGlnAlaAlaArgLeuAlaAspThrAlaAsnTyrThr 220
 DB 601 ACGGGGAGCAGACAGCTAGTCGTGCGTCAGCGCCGCTGGCCGACACGGCCCACTACACC 660

QY 221 CysValAlaIysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240
 DB 661 TGTGTGCCCAAGAACATCGTAGCCCGTCGCCAAGCACCTCTGCAGCGGTCAITGTTTAT 720

QY 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
 DB 721 GTGAACGGTGGGTGGTGCAGCTGACCTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGCGT 780

QY 261 GlyTrpGlnIysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
 DB 781 GGTCTGGCAGAAACGGAGCCGAGCTGCACCAACCCGCGACCTCTCAACGGGGGCGCCTTC 840

QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
 DB 841 TGTGAGGGGCGAATATGTC---CAGAAACAGCCTGCGCCACTCTGTGTGCCAGTGGATGGG 897

QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
 DB 898 AGCTGGAGTTCTGGAGTAAGTGGTGCAGCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
 DB 958 CGCGAGTGTCTGTACCCAGCACCCCGCAATGGAGGTGAGGAGTGTGGGGTGTGCTGACCTG 1017

QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
 DB 1018 GACCCCGCAACTGTACCAAGTACCTGTGCTGCACACCGCTCTTGTGCCCGGAGGAGCTG 1077

QY 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
 DB 1078 GCTCTCTACATCGCCCTTGTGCTGTGGCTGTGTGCTCTTCTTGTGTGTGGCCCTT 1137

QY 381 IleLeuValTyrCysArgIysLysGluGlyLeuAspSerValAlaAspSerSerIle 400
 DB 1138 GGACTCATTTACTGTGCGAAGAGGAGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerIysAlaAspAsnProHisLeu 420
 DB 1198 CTCACTTCGGGCTTCAGGCTGTGACATCAAGCCCGAGCAAGAGCAGCAACCCCACTG 1257

QY 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThrThr 439
 DB 1258 CTCACCATCCAGCCAGACCTCAGCACCACTACCACTACCAAGGAGGAGCTATGTTCG 1317

QY 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
 DB 1318 AGGCAGATGGAGCCAGCCCAAGTTCCAGCTCTCTAATGGTCACCTGTGCTCAGCCCACTG 1377

QY 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluPheVal 479
 DB 1378 GGAAGTGGCGGCGCATACGTTGCACACAGCTACCCACCTCTGAGGCTGAGGACTTCGTC 1437

QY 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
 DB 1438 TCCCGCTCTCCACCCCAAAACTACTTTTGTTCCTCCCGCGGCGCACCAAGCAACATGGCC 1497

QY 500 TyrGlyThrPheAsnProGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
DB 1498 TACGGGACCTTCAACTCTTCGGGGGGCGGCTGATGATCCTTAATACGGGATCAGCCCTC 1557
QY 520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
DB 1558 CTATACCCCGGATGCCATCCCGGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617
QY 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
DB 1618 AAGCCAGAACGCTGAGGTTGCCCTAGCTGGCTGCAGACCTCTGCTGAGTCCAGTCGTT 1677
QY 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
DB 1678 AGCTGTGGGCCCCAGAGTCTCTCACCGGCAGTCACTCTTGCATGGACCACTGT 1737
QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
DB 1738 GGAGAGCCAGCCCTGACAGCTGGAGTCTGGCTCAAAAGCAGTCTCGAGGGCAGT 1797
QY 600 TrpGluGlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGln 619
DB 1798 TGGGAG---GATGTCTGCACCTTGTGTGAGGAGTCACTTCCCACTCTACTACTGCAG 1854
QY 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
DB 1855 CTGGAGCGCGGGCTCTGTGTCTTACGAGGAGCTGGGCGCTTGTCCCTCGTAGGA 1914
QY 640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla 659
DB 1915 GAGCCCTCAGCGTGGCTGCCACCAAGCGCTCAGGCTCTCTTGTCTCCGCTGGCC 1974
QY 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
DB 1975 TGTAGTCCCTTGAGTACAACTCGAGTACTGCTCTACAGCAGACCCAGCAGCTCTC 2034
QY 680 LysGluValValGlnLeuGlyGlnLeuGlyGlnLeuIleGlnGluProArgVal 699
DB 2035 AAGGAGTGGTGGAGTGGAGAGCAGTGGTGGAGAGTGGTGGAGAGCTTCGGCTC 2094
QY 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
DB 2095 CTGCACCTCAAGACAGTATACCAACCTAGTCTCTCCATCCAGCAGTCCCGCAGCTCC 2154
QY 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
DB 2155 CTGTGGAGCAAGCAAGTACTTGTGCTACAGTACAGGAGATCCCTTTTACCATCTGGAAC 2214
QY 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
DB 2215 GGCACCCAGCAGTATCTGACCTGCACCTTCCCTGGAGCGCATCAACGCCAGCACCCAGC 2274
QY 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
DB 2275 GACCTGGCTGCAAGGTGGTGTGGTGGAGTGGAGGAGATGGCGAGCTTCAACATC 2334
QY 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuAlaLeuGluSerGluAla 799
DB 2335 AACCTCAACATCACTAAGGACACAGGTTGTGTAATTGTGGCTCTGGAGAGTGAAGGG 2394
QY 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
DB 2395 GGGTTCACGACCTGGTGGGCCCCAGTGGCTTCAGATCCCTTCTCATTCGGCAAG 2454
QY 820 IleLeuSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
DB 2455 ATCATCGCAGCTGGACCCACCTCTGCAGCGGGCGCGCATGTGAGAACTCTAGCCAG 2514
QY 840 LysLeuHisLeuAspSerHisLeuSerPheAlaSerLysProSerProThrAlaMet 859
DB 2515 AAATTCACCTGGAGCAGCATCTTGTCTTGTCTTCAAGCCCGCCCTACAGCCATG 2574
QY 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879

DB 2575 ATCTCAACCTATGGGAGGACGCGCACTTCCCAACGCGCAACCTCGGCCAGCTGGCAGCA 2634
QY 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
DB 2635 GCTGTGCGCGAGTGGGCGCAACAGATGCTGGCTCTTTCACGGTGTGGAGCGCGAGTGT 2694
RESULT 6
AAK52261
ID AAK52261 standard; cDNA; 2907 BP.
XX
AC AAK52261;
XX
XX 06-NOV-2001 (first entry)
XX Human polynucleotide SEQ ID NO 806.
DE
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
XX Homo sapiens.
XX
XX WO200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US004098.
XX
XX 03-FEB-2000; 2000US-00496914.
XX
XX 27-APR-2000; 2000US-00560875.
XX
XX 20-JUN-2000; 2000US-00598075.
XX
XX 19-JUL-2000; 2000US-00620325.
XX
XX 01-SEP-2000; 2000US-00654936.
XX
XX 15-SEP-2000; 2000US-00663561.
XX
XX 20-OCT-2000; 2000US-00693325.
XX
XX 30-NOV-2000; 2000US-00728422.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPT; 2001-476283/51.
XX
XX P-PSDB; AAM79128.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
in diagnosis and gene therapy.
XX
XX Claim 1; Page 2691-2694; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
XX Sequence 2907 BP; 552 A; 966 C; 881 G; 508 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 1,84e-296 Length: 2907
Score: 4434.00 Matches: 853
Percent Similarity: 90.78% Conservative: 4

Best Local Similarity: 90.36% Mismatches: 17
 Query Match: 92.63% Indels: 70
 DB: 4 Gaps: 7

US-09-970-944-2 (1-899) x AAKS2261 (1-2907)

QY	1	MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu	20
DB	148	ATGACGAGGGTCCCTCCCTG-----ATGGCGGGCAGACAG	183
QY	21	ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro	40
DB	184	CACGGATGGAGCGCCAGCAGAGTGCACCGTGGCCAAACAGTGCCTGGTGCACACCG	243
QY	41	AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro	60
DB	244	GACCTGCTTCCCACTTCCCTGGTGGAGCCGAGGATGTACATCGTCAAGAACAAAGCCA	303
QY	61	ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu	80
DB	304	GTGCTGCTTGTGTGCAAGGCGCGTCCCGCCACGACAGATCTTCTCAAGTGCACGGGGAG	363
QY	81	TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro	100
DB	364	TGGGTGGCGCAGGTGGACACAGTATCGAGCGCAGCAGACAGCGGAGCAGTGGGCTGCC	423
QY	101	ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu	120
DB	424	ACCATGAGGTCCGATATATGCTCAAGGAGCAGGTTCGAGAGGTTCGGGGCTGGAG	483
QY	121	GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla	140
DB	484	GAATACTGTGTCAGTGGTGGATGAGTCTCTCGGCGCACCCACCAAGAGTCAAGAGGCC	543
QY	141	TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal	160
DB	544	TACATCCGCATAGCCTATTTCGCAAGAACTTCGAGCAGGAGCGCTGGCCAAAGAGGTG	603
QY	161	SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu	180
DB	604	TCCTTGAGCAGGAGCATCGTGTGCTGCCCTGCTCCACCGAGGCGCATCCCTCAGCGCGAG	663
QY	181	ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle	200
DB	664	GTGAGTGGCTCCGGAACAGGAGCCTGGTGGACCCCGTCCCTGGACCCCATGTATACATC	723
QY	201	ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr	220
DB	724	ACGCGGAGCACAGCCTGGTGGTGGACAGGCGCGCTTGCACAGCGGCCCTTGCACAGCGCAACTACACC	783
QY	221	CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr	240
DB	784	TGCGTGGCCNAGAACATCGTGGCAGCTGCGCGCAGCGCCTCCCGTGTGTATCTGCTAC	843
QY	240	-----240	
DB	844	GGTGGGCGCGGACTCCCTGTTGCACAGGAGAGGCATCGCGTGGCCCTGGCAGTGAC	903
QY	241	-----ValAsnGlyClyTrpSerThrTrpThrGluTrp	251
DB	904	ATGTGCTGTCTCTGTGTCGGCCAGTGAACGGTGGGTGGTGGAGCCGAGTGG	963
QY	252	SerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsn	271
DB	964	TCCGTCTGCAGCCAGCTGTGGGCGCGCTGGCAGAAACAGGAGCGGAGCTGCACCAAC	1023
QY	272	ProAlaProLeuAsnGlyAlaPheCysGluGlyClnAsnValHisAspArgThrVal	291
DB	1024	CCGGCGCTCTCAACGGGGCGCTTCTGTGAGGGCGAGATGTC---CAGAAAAAGGCC	1080
QY	292	SerSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCys	311
DB	1081	TGCGCCACCTGTGCCAGTGGAGCGGAGCTGGAGCCCGTGGAGCAAGTGTTCGGCTGT	1140

QY	312	GlyLeuAspCysThrHisTrpArgSerArgGluCysSerAspProAlaProArgAsnGly	331
DB	1141	GGGCTGAGCTGCACCCACTGGCGAGCCGTGAGTGCTCTGACCCAGCAGCCCGAACGGA	1200
QY	332	GlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysVal	351
DB	1201	GGGAGAGTGCACAGGCATGACCTGGACACCCGCAACTGTACAGTGACCTCTGTGTA	1260
QY	352	His-----SerAlaSerGly	356
DB	1261	CACAACCTCTACACCCCTGCCCCCACCAGGCGCATGCTGCTCCCGCAGCTGCTTCTGC	1320
QY	357	ProGluAspValAlaLeuTyrValGlyLeuAlaValAlaValCysLeuValLeuLeu	376
DB	1321	CCTGAGACGTGGCCCTCTATGTGGCCCTCATCGCGTGGCGCTGCTGCTGCTGCTG	1380
QY	377	LeuLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAla	396
DB	1381	CTGCTTGTCT	1440
QY	397	AspSerSerIleLeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAsp	416
DB	1441	GACTCGTCCATTCTCCTCAGCTTCAGCCCGTTCAGCATCAAGCCAGCAGCAGCAGCAG	1500
QY	417	AsnProHisLeuLeuThrIleGlnProAspLeuSer---ThrThrThrThrThrGlnGly	435
DB	1501	AAACCCCATCTGCTCACCATCCAGCCCGGACCTCAGCACCCACCCACCTACCCAGGC	1560
QY	436	SerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeu	455
DB	1561	AGTCTCTGTCCTCCCGGAGGATGGGCCAGCCCAAGTTCAGCTCACCATGGGACCTG	1620
QY	456	LeuSerProLeuGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAla	475
DB	1621	CTCAGCCCTGCTGGTGGCGCGCCACACACTGCACCCAGCTCTCCCTCTGAGGCC	1680
QY	476	GluGluPheValSerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThr	495
DB	1681	GAGAGTGTCTCTCCCGCTCTCCAGCCAGAACTACTTCCGCTCCCTGCGCCGAGGCACC	1740
QY	496	SerAsnMetThrTyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThr	515
DB	1741	AGCAACATGACCTATGGGACCTTCAACTTCTCGGGGCGCGCTGATGATCCCTAATACA	1800
QY	516	GlyIleSerLeuLeuLeuProAspAlaIleProArgGlyLysIleTyrGluIleTyr	535
DB	1801	GGAAATCAGCTCTCTATCCCGCAGATGCATACCCCGAGGGAAGATCTATGAGATCTAC	1860
QY	536	LeuThrLeuHisLysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeu	555
DB	1861	CTCAGCTGCACAAAGCCGGAAGAC-----1884	
QY	556	SerProIleValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAla	575
DB	1885	-----GTGAGCTGTGGACCCCTGCGTCTGCTCACCAGCCAGTATCTCTGGCT	1935
QY	576	MetAspHisCysGlyGluProSerProAspSerTrpSerLeuArgLysLysGlnSer	595
DB	1936	ATGGACCATGTGGGAGCCAGCCCTGCAGCTGAGCGCTGCGCTCAAAAAGCAGTCG	1995
QY	596	CysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeu	615
DB	1996	TGCGAGGGAGCTGGAG---GATGTGCTGCACCTGGGCGAGGAGCGCCCTCCACCTC	2052
QY	616	TyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPhe	635
DB	2053	TACTACTGCAGCTGGAGCCAGTGCCTGCTAGCTTTCACGAGCAGCTGGGCGCCCTT	2112
QY	636	AlaLeuValGlyGluAlaLeuSerValAlaAlaLysArgLysLeuLeuLeuPhe	655
DB	2113	GCCCTGGTGGAGAGGCCCTCAGCGTGGCTGCGCCGCAAGCCCTCAAGCTCTCTGTTT	2172

```
QY 656 AlaproValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThr 675
Db 2173 GCGCCGGTGGCCCTGCACCTCCCTCGAGTACAATCCGGGTCTACTCCCTGCATGACACC 2232
QY 676 HisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGln 695
Db 2233 CACGATGCATCAAGGAGGTGGTGCAGTCGAGAGACAGCTGGGGGACAGCTGATCCAG 2292
QY 696 GluProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAsp 715
Db 2293 GAGCCAGGGTCTCTGCACCTTCAAGGACGATGACACAACTGGCCCTATCCATCCAGAT 2352
QY 716 ValProSerSerLeuTyrLysSerLysLeuValSerTyrGlnGluIleProPheTyr 735
Db 2353 GTGCCAGCTCCCTGTGGAGAGTAAGCTCTTGTCTAGCTACCAAGGAGATCCCTTTTAT 2412
QY 736 HisIleTyrAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSer 755
Db 2413 CACATCTGGAATGGCAGCAGCGGTACTTGCACCTGCACCTTACCCCTGGAGCGTGCAGC 2472
QY 756 ProSerThrSerAspLeuAlaCysLysLeuTyrValTyrGlnValGluGlyAspGlyGln 775
Db 2473 CCCAGCAGTACTAGTACCTGGCTGCAGCTGTGGTGTGGCAGGTGGAGCGGCGGCGAG 2532
QY 776 SerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuAlaLeu 795
Db 2533 AGCTTCAGATCAACTTCAACATCACCAGAGACAAAGGTTTGTGAGCTGCTGGCTCTG 2592
QY 796 GluSerGluAlaGlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeu 815
Db 2593 GAGAGTGAAGCGGGGTCCAGCGCTGTGGGCCAGTGCCTTCAAGATCCCTTCTCTC 2652
QY 816 IleArgGlnLysIleLeuSerSerLeuAspProCysArgArgGlyAlaAspTyrArg 835
Db 2653 ATTGGCAGAGATAATTTCCAGCTCGACCCCTGTAGCGGGGTGGCGACTGGCGG 2712
QY 836 ThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSer 855
Db 2713 ACTGTGCCCGCAGAACTCCACCTGGACAGCATCTCAGCTTCTTGTCTCCAGCCGAGC 2772
QY 856 ProThrAlaMetIleLeuAsnLeuTyrGluAlaArgHisPheProAsnGlyAsnLeuSer 875
Db 2773 CCCACAGCCATGATCTCTCAACCTGTGGAGCGCGGCACCTTCCCAACGGCAACCTCAGC 2832
QY 876 GlnLeuAlaAlaValAlaGlyLysGlnProAspAlaGlyLeuPheThrValSer 895
Db 2833 CAGCTGGCTCAGCAGTGGTGGACTGGCGGAGCCAGCAGCGTGGCCCTTTCACAGTGTGC 2892
QY 896 GluAlaGluCys 899
Db 2893 GAGGCTGAGTGC 2904
RESULT 7
ABK15169
ID ABK15169 standard; DNA; 3580 BP.
XX
AC ABK15169;
XX
DT 23-APR-2002 (first entry)
XX
DE Human REPTR 1 cDNA sequence.
XX
KW REPTR; human; antiinflammatory; cytostatic; immunosuppressive; antiviral;
KW anti-HIV; antiarthritic; anticonvulsant; neurotropic; neuroprotective;
KW antiallergic; antibody; immunogen; endometriosis;
KW gastrointestinal disorder; gastritis; oesophageal carcinoma;
KW Crohn's disease; irritable bowel syndrome; ulcerative colitis;
KW endocrine disorder; hypothalamus disorder; Kallman's disease;
KW autoimmune disease; inflammatory disease; infertility; receptor;
KW acquired immune deficiency syndrome; AIDS; rheumatoid arthritis; allergy;
KW osteoarthritis; diabetes mellitus; multiple sclerosis;
KW systemic lupus erythematosus; cell proliferative disorder; cancer;
KW developmental disorder; Duchenne muscular dystrophy; gene;
```

```
KW Becker muscular dystrophy; neurological disorder; epilepsy;
KW Alzheimer's disease; Huntington's disease; reproductive disorder; ss.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 4..2532
XX /tag= a
XX /product= "REPTR1 protein"
XX
XX WO200198354-A2.
XX
XX 27-DEC-2001.
XX
XX 21-JUN-2001; 2001WO-US019942.
XX
XX 21-JUN-2000; 2000US-0214027P.
XX 25-AUG-2000; 2000US-0228045P.
XX 12-DEC-2000; 2000US-0255104P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Griffin JA, Kallick DA, Tribouley CM, Yue H, Nguyen DB, Tang YT;
XX Lal P, Policky JL, Azimzai Y, Lu DAM, Graul R, Yao MG, Burford N;
XX Hafalia AJA, Baughn MR, Bandman O, Patterson C, Yang J, Xu Y;
XX Gandhi AR, Warren BA, Ding L, Sanjanwala MS, Duggan BM, Lu Y;
XX
XX WPI; 2002-090432/12.
XX P-PSDB; AAU17818.
XX
XX Twelve human receptors (referred to as REPTR-1 to REPTR-12), useful in
XX the diagnosis, treatment and prevention of gastrointestinal (e.g.
XX gastritis), autoimmune/inflammatory (e.g. osteoarthritis) and cell
XX proliferative (e.g. cancer) disorders.
XX
XX Claim 57; Page 142-143; 157pp; English.
XX
XX This invention relates to twelve human receptors cDNA sequences referred
XX to as REPTR-1 to REPTR-12), and the proteins encoded thereby. The
XX proteins of the invention may have antiinflammatory, cytostatic, active
XX immunosuppressive, antiviral, anti-HIV, antiarthritic, muscular active
XX general, anticonvulsant, neurotropic, neuroprotective, antiallergic
XX activities. The sequences of the invention may be used to produce REPTR
XX agonists or antagonists, and the protein sequences may be used to raise
XX anti-REPTR antibodies. These molecules and the REPTR polynucleotides and
XX polypeptides of the invention are useful in the diagnosis, treatment and
XX prevention of gastrointestinal (e.g. gastritis, oesophageal carcinoma,
XX Crohn's disease, irritable bowel syndrome, ulcerative colitis), endocrine
XX (e.g. hypothalamus disorder, Kallman's disease), autoimmune/ inflammatory
XX (e.g. acquired immune deficiency syndrome (AIDS), rheumatoid arthritis,
XX allergies, osteoarthritis, diabetes mellitus, multiple sclerosis,
XX systemic lupus erythematosus), cell proliferative (e.g. cancer),
XX developmental (e.g. Duchenne and Becker muscular dystrophy), neurological
XX (e.g. epilepsy, Alzheimer's disease, Huntington's disease) and
XX reproductive (e.g. infertility, endometriosis) disorders. Numerous other
XX examples of each disorder are given in the specification. The present
XX sequence represents the human REPTR1 cDNA sequence of the invention
XX
XX SQ Sequence 3580 BP; 670 A; 1233 C; 1025 G; 652 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.:	1.41e-292	Length:	3580
Score:	4379.50	Matches:	838
Percent Similarity:	93.22%	Conservative:	1
Best Local Similarity:	93.11%	Mismatches:	2
Query Match:	91.49%	Indels:	59
DB:	6	Gaps:	3

US-09-970-944-2 (1-899) x ABK15169 (1-3580)

Qy	1	MetAlaValArgProGlyLeuTyrProAlaLeuLeuGlyIleValLeuAlaAlaTyrLeu	20
Db	4	ATGGCCGCTCCGGCCGCGCTGTGGCCAGCGCTCTCTGGGCATAGTCTCGCGCTTGCTC	63

QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
DB 64 CGCGGCTCGGGTGCCAGCAGAGTGCACCGTGGCCAAACCAGTGGCTGGTGCAACCCG 123
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
DB 124 GACCTGCTTCCCCACTTCCTTGGTGGAGCCGAGAGTGTACATCGTCAAGAACAGCCA 183
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
DB 184 GTCTGCTGTGTGCAAGGCCGTGCCCGCCACGACAGATCTTCTCAAGTGCACAGGGGAG 243
QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
DB 244 TGGGTGGCCAGGTGGACCATGATCGAGCGAGCGACACAGCGGAGCATGGGCTGCC 303
QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
DB 304 ACCATGGAGTCCGCATTAATGTCTCAAGGCAGCAGGTTCGAGAAGGTGTTTCGGGCTGGAG 363
QY 121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
DB 364 GAATATCGGTGCGAGTGGCATGGAGCTCTCGGGCACCAACCAAGAGTCAAGAGGCC 423
QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
DB 424 TACATCCGATACCTATTGTCGCAAGACTTCGAGCAGGAGCGCTGGCCAAAGAGGTG 483
QY 161 SerLeuGlnGlnGlyIleValLeuProCysArgProProGluGlyIleProAlaGlu 180
DB 484 TCCTCGAGCAGGGCATCGTGTCTGCCCTGCCGTCCACCGAGGGCATCCCTCCAGCCGAG 543
QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerSerLeuAspProAsnValTyrIle 200
DB 544 GTGGAGTGGCTCCGGAACAGAGACCTGGTGGACCCCGTCCCTGGACCCCAATGTATACATC 603
QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
DB 604 ACSCGGGAGCACAGCTGTGTGTCGACAGCCGCGCTTCTGACAGCGCCAACTACACC 663
QY 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240
DB 664 TGGTGGCCCAAGAAATCGTGGCAGCTCGCGCAGCGCCTCCGCTGCTCATCGTCTAC 723
QY 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
DB 723 ----- 723
QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
DB 723 ----- 723
QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
DB 724 -----GTGGACGGC 732
QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
DB 733 AGCTGAGCCCGTGGAGCAAGTGGTGGCTGTGGCTGGACTGGACCCCACTGGCGGAGC 792
QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
DB 793 CGTGAAGTCTCTGACCCAGCACCCCGAACGAGGGGAGAGTGGCCAGGCACTGACCTG 852
QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
DB 853 GACACCCGCAACTGTACCACTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTG 912
QY 361 AlaLeuTyrValGlyIleuIleAlaValAlaValCysLeuValLeuLeuValLeu 380
DB 913 GCCCTCATGTGGGCGCTCATCGCGGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCT 972

QY 381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
DB 973 ATCTCTGTTTATTTGCCGGAAGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1032
QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
DB 1033 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAAGCAAGCAGACACCCCATCTG 1092
QY 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
DB 1093 CTCACCATCCAGCCGAGCTCAGCACCAACCAACCACTTACCAGGGCAGTCTCTGTCCC 1152
QY 440 ArgGluAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
DB 1153 CGGACAGATGGGCCAGCCCCAAGTTCAGCTCACAATGGGCACCTGTCTAGCCCCCTG 1212
QY 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
DB 1213 GGTGGCGGCGCCACACACTGCACCACTCTCCACCTCTGAGCCGAGGAGTTCGTC 1272
QY 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
DB 1273 TCCCGCTCTCCACCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAACATGAC 1332
QY 500 TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
DB 1333 TATGGACCTTCACTTCTCGGGGCGGCTGATGATCCCTAATACAGGAATCAGCCT 1392
QY 520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
DB 1393 CTCATCCCCCAGATGCCATACCCGAGGAGAGATCTATGAGATCTACCTACGCTGCAC 1452
QY 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
DB 1453 AAGCCGGAAGACGTAGGTGGCTTGGCTGTGTCAGACCTGTGATGTCCTCATCGTT 1512
QY 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
DB 1513 AGCTGTGGACCCCTGGCGTCTGCTACCCGCGCCAGTCATCTCTGCTATGGACCACTGT 1572
QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
DB 1573 GGGGAGCCCGACCTCAGACCTGGAGCTGGCTCAAAAAGCAGTCTGCGAGGGCAGC 1632
QY 600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
DB 1633 TGGGAG---GATGTGTGCACCTGGCGGAGGAGCGCCCTCCACCTCTACTACTGCCAG 1689
QY 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
DB 1690 CTGGAGCCAGTGCCTGTCTACGTCTTACCGAGCAGCTGGCGCGCTTGGCCCTGGTGGGA 1749
QY 640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla 659
DB 1750 GAGGCCCTCAGCGTGGCTGCCCGCAAGCGCTCAAGCTGCTTCTGTTCGCCCGGTGGCC 1809
QY 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
DB 1810 TGCACCTCCCTCGAGTACAACATCCGGTCTACTGCTTCATGACACCAACCACTGCACTC 1869
QY 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
DB 1870 AAGGAGTGGTGCAGCTGGAGAGCAGCTGGGGGAGACAGCTGATCCAGAGGCCACGGTTC 1929
QY 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
DB 1930 CTGCATCTCAAGACAGTATACCAACCTGGCCCTATCCATCCACGATGTGCCAGCTCC 1989
QY 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
DB 1990 CTGTGAAGAGTAAGCTCCTTGTCTGCTACCTACAGGAGATCCCTTTTATCACATCTGAA 2049
QY 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759

```

Db      2050 GGCACGAGCGGACTTGGCACTTGCACCTTCCCTGGAGCGTGTACGCCCCAGACCTAGT 2109
QY      760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db      2110 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGTGGAGGGCGACGGGACAGAGCTTCAGCATC 2169
QY      780 AspPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db      2170 AACTTCAACATCAACAGGACACAAAGGTTTGTCTGAGCTGTGGCTCTGGAGAGTGAAGCG 2229
QY      800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db      2230 GGGGTCCACGCTGTGGGCCCCAGTTCCTTCAAGATCCCTTCTCATTCGGCAGAG 2289
QY      820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db      2290 ATAATTTCCAGCTGGACCCACCTGTAGGCGGGGTGCCGACTTGGCGACTCTGGCCAG 2349
QY      840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db      2350 AAATCCACCTGCAGACCCATCAGCTTCTTGGCTCCAAAGCCCGACGCCACAGCCATG 2409
QY      860 IleIleAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db      2410 ATCTCAACTGTGGAGGGCGGCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2469
QY      880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db      2470 GCAGTGGCTGGACTGGCCAGCCAGACGCTGGCTTTCACAGTGTGGAGGCTGAGTGC 2529

RESULT 8
AAV52941
ID      AAV52941 standard; cDNA; 1787 BP.
XX      AC
XX      AAV52941;
XX      25-MAR-2003 (revised)
XX      21-DEC-1998 (first entry)
XX      Human UNC-5 homologue unc5h-1 cDNA.
XX      DE
XX      UNK-5; UNC5H-1; human; netrin receptor; cell migration; axon guidance;
XX      diagnosis; therapy; ds.
XX      OS
XX      Homo sapiens.
XX      FH
XX      CDS
XX      3..1670
XX      Location/Qualifiers
XX      /*tag= a
XX      /transl_except= (pos:21..22, aa:Xaa)
XX      /transl_except= (pos:200..203, aa:Ile)
XX      /transl_except= (pos:771..772, aa:Ser)
XX      /transl_except= (pos:785..786, aa:Leu)
XX      /transl_except= (pos:1078..1079, aa:Xaa)
XX      /transl_except= (pos:1098..1099, aa:Xaa)
XX      /transl_except= (pos:1106..1107, aa:Xaa)
XX      /transl_except= (pos:1621..1622, aa:Gly)
XX      /note= "these codons have apparent 1 nucleotide deletions
XX      or insertions, which alter the reading frame"
XX      PN
XX      WO9837085-A1.
XX      PD
XX      27-AUG-1998.
XX      PF
XX      19-FEB-1998; 98WO-US003143.
XX      PR
XX      19-FEB-1997; 97US-00808982.
XX      PA
XX      (REGC ) UNIV CALIFORNIA.
XX      PI
XX      Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasu K;

```

```

DR      WPI: 1998-495364/42.
DR      P-PSDB; AAV78899.
XX      Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and
XX      the biopharmaceutical industry.
XX      Claim 7; Page 17; 32pp; English.
XX      This partial cDNA, termed unc5h-1, comprises a human homologue of
XX      Caenorhabditis elegans unc-5, unc5h-1 and unc5h-2 (see AAV52943) cDNAs
XX      were isolated from an embryonic brain cDNA library. The predicted
XX      proteins (see AAV78899 and AAV78901) show similarity with UNC-5. They are
XX      predicted to be involved in cell migration and axon guidance, and are
XX      characterised as receptor proteins for netrins. Gene expression is
XX      observed in regions where differentiating neurons are undergoing
XX      axogenesis. Rat unc5h-1 (see AAV52940) and unc5h-2 (see AAV52942) cDNAs
XX      are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly
XX      from transfected host cells by utilising these vertebrate UNC-5 nucleic
XX      acids. The invention also provides unc-5 hybridisation probes and
XX      primers, vertebrate UNC-5-specific binding agents such as specific
XX      antibodies, and methods of making and using the subject compositions in
XX      diagnosis (e.g. genetic hybridisation screens for vertebrate unc-5
XX      transcripts), therapy (e.g. gene therapy to modulate vertebrate unc-5
XX      gene expression) and in the biopharmaceutical industry (e.g. as
XX      immunogens, reagents for modulating cell guidance, reagents for screening
XX      chemical libraries for lead pharmacological agents, etc.). (Updated on 25
XX      -MAR-2003 to correct PI field.)
XX      SQ      Sequence 1787 BP; 349 A; 603 C; 491 G; 344 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      156-176      Length:      1787
Score:          2694.00      Matches:    543
Percent Similarity: 97.14%      Conservative: 1
Best Local Similarity: 96.96%      Mismatches: 11
Query Match:    56.28%      Indels:     11
DB:             2           Gaps:       2

US-09-970-944-2 (1-899) x AAV52941 (1-1787)
QY      344 AsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363
Db      3 AACTGTACCACTGACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTCCTCTAT 61
QY      364 ValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeuVal 383
Db      62 GTGGGCTCATCGCCGTGGCCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
QY      384 TyrCysArgLysLysGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSer 403
Db      122 TATTGCCGGAAGAGGAGGGCTGGACTCAGATGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCT 181
QY      404 GlyPheGlnProValSerIle-LysProSerLysAlaAspAsnProHisLeuLeuThrIle 423
Db      182 GGCTTCCAGCCCTCAGCATCTTAAGCCAGCAAAAGCAGACAAACCCCATCTGCTCCACAT 241
QY      423 eGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThrThrThrThrThrThr 442
Db      242 CCAGCCGAGCTCAGACACACACACACACACACACACACACACACACACACACACACACAC 301
QY      442 pGlyProSerProLysPheGlnLeuThrAenGlyHisLeuLeuSerProLeuGlyGlyGly 462
Db      302 TGGGCCCGAGCCCAAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGT 361
QY      462 YArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSerArgLe 482
Db      362 CCGCCACACACTGCACACAGCTCTCCACCTCTGAGGCCGAGAGTCTGCTCTCCGCCCT 421
QY      482 uSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrThrGlyTh 502
Db      422 CTCACCCAGAACTACTTCCGCTCCCTCCGCCGAGGACCCAGCAACATGACCTATGGGAC 481
QY      502 rPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePr 522

```

Db 482 CTTCAACTCTCGGGGCGGCTGATGATCCCTAATACAGGATCAGCCTCTCATCCC 541
 QY 522 oProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGI 542
 Db 542 CCAGATGCGCATACCCGAGGAGATCTATGAGATCTACCTCAGCTGCACAGCGGA 601
 QY 542 uAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGI 562
 Db 602 AGAGTGGAGTTGCCCTAGCTGGCTGTGACAGCCCTGCTGAGTCCCATCTAGCTGTGG 661
 QY 562 YProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluPr 582
 Db 662 ACCCCTGGCGTCTCTCACCAGGAGATCATCTGGCTATGACCACTGTGGGAGCC 721
 QY 582 oSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGI 602
 Db 722 CAGCCTTGACAGCTGGAGCGCTGCCCTCAAAAAGCAGTCTGCGAGGAGC-TGGGAG-- 778
 QY 602 nAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAl 622
 Db 779 -GATGT-CTGCACCTGGCGAGAGGCGCCCTCCACCTCTACTGCGCAGCTGGAGGC 836
 QY 622 aSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLe 642
 Db 837 CAGTGCCTGCTACGTCTTACCAGCAGCTGGCGCGCTTTGGCTGTGTGGAGAGCCCT 896
 QY 642 uSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAlaCysThrSe 662
 Db 897 CAGCGTGGCTGCCGCGCAAGCGCTCAAGCTGCTTCTGTTTGGCGCGGTGGCCCTGCACCTC 956
 QY 662 rLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluVa 682
 Db 957 CCTCGAGTCAACATCCGGGTCTACTGCTGCTGATGACCCAGCATGCATCAAGGAGGT 1016
 QY 682 lValGlnLeuGluLysGlnLeuGlyGlnLeuIleGlnGluProArgValLeuHisPh 702
 Db 1017 GGTGACGTGGAGAGCAGCTGGGGGACAGCTGATCCAGAGCCACGGGTCTGCACCT 1076
 QY 702 eLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLy 722
 Db 1077 -AAGGACAGTTACCAACACCTGCC-CTATC-ATCCACAGATGTCGCCAGCTCCCTGTGGAA 1133
 QY 722 sSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGI 742
 Db 1134 GAGTAGCTCCTTGTAGCTACAGGAGATCCCTTTATCATCTGGAGTGGACGCA 1193
 QY 742 nArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAl 762
 Db 1194 GCGGTACTTGACCTGCACCTTCAACCTGGAGCGGTGTCAGCCCGCAGCACTAGTCACTGGC 1253
 QY 762 aCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAs 782
 Db 1254 CTGCAAGCTGTGGGTGTGGCAGTGGAGGCGGACGGGAGAGCTTCAGCATCAACTTCAA 1313
 QY 782 nIleThrLysAspThrArgPheAlaGluLeuAlaLeuGluSerGluAlaGlyValPr 802
 Db 1314 CATCACCAAGGACACAGGTTTCTGAGCTGCTGAGAGAGTGAAGCGGGGTCCC 1373
 QY 802 oAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleSe 822
 Db 1374 AGCCCTGGTGGGCGCCAGTGCCTTCAAGATCCCTTCTCTTCCGCGAGAGATATTTC 1433
 QY 822 rSerLeuAspProProCysArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHi 842
 Db 1434 CAGCCTGGAGCCACCTCTAGGCGGGGTGCCGACTGCGGAGCTCGCCCGAGAACTCCA 1493
 QY 842 sLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAs 862
 Db 1494 CTTGGACAGCCATCTCAGCTTCTTTGCTCCAGCCAGCCCGCCACAGCCATGATCTCAA 1553
 QY 862 nLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAl 882

Db 1554 CCTGTGGAGGCGGCGCACTTCCCAACGCAACCTCAGCCAGCTGCTGCAGCAGTGGC 1613
 QY 882 aGlyLeuGlyGlnProAspAlaGlyLeu-PheThrVal-SerGluAlaGluCys 899
 Db 1614 TGGGACTGGCCAGCAGGACGGTGGCTTCTTTCAAGTGTTCGGAGGCTGAGTGC 1667
 RESULT 9
 ID AAS75738 standard; cDNA; 3646 BP.
 XX AAS75738;
 XX 13-FEB-2002 (first entry)
 XX DNA encoding novel human diagnostic protein #11542.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX Homo sapiens.
 XX WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US008631.
 XX 31-MAR-2000; 2000US-00540217.
 XX 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 XX P-PSDB; ABG11551.
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX Claim 1; SEQ ID NO 11542; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have application of mutations
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 3646 BP; 930 A; 917 C; 921 G; 877 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 1e-174 Length: 3646
 Score: 2673.50 Matches: 507
 Percent Similarity: 72.40% Conservative: 154
 Best Local Similarity: 55.53% Mismatches: 221

Query Match:	55.85%	Indels:	31
DB:	5	Gaps:	12
US-09-970-944-2 (1-899) x AAS75738 (1-3646)			
QY	9	ProAlaLeuLeuGlyIleValLeuAlaAtrpLeuArgGlySerGlyAlaGlnGlnSer	28
DB	230	CCTGGCCCTG-----GCCCTGCTCAGCGCCAGCGGCACTGGCTCCGCGCCCAAGATGAT	283
QY	29	-----AlaThrValAlaAenProValProGlyAlaAasnProLeuLeuLeuProHis	45
DB	284	GACTTTTTTCATGAACCTCCCAAGAACTTTTCCTTCCTGATCCACCTGAGCCCTCTGCCACAT	343
QY	46	PheLeuValGluProGluAspValTyrIleValIysAsnIysProValLeuValCys	65
DB	344	TTCTTATTGAGCCTGAAGAGCTTATATTGTGAAGAAATAGCTTCCGCTCTCATTTGCCGGAGTGACTGT	403
QY	66	LysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArgGlnVal	85
DB	404	AAAGCAAGCCCTGCCACCCAGATCTATTTCAGTGTAAATAGTGAATGGGTTTCATCAGAAG	463
QY	86	AspHisValIleGluArgSerThrAspGlySerSerGlyGluProThrMetGluValArg	105
DB	464	GACCACATAGTAGATAAAGAGTAGATGAACCTTCCGCTCTCATTTGCCGGAGTGAGC	523
QY	106	IleAsnValSerArgGlnGlnValGluIysValPheGlyLeuGluGluTyrTrpCysGln	125
DB	524	ATTGAGATTTCGCGCCAGCAAGTGAAGAACTCTTTGGACCTGAAGATTACTGGTGCAG	583
QY	126	CysValAlaTrpSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAla	145
DB	584	TGTGTGCGCTGGAGCTCCGCGGTACCACAAAGACCGGAGGCGTATGTGGCAATTGCA	643
QY	146	ArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGluGlnGly	165
DB	644	TATCTACGGAAGACATTGTGACGAGAACCCCTAGGAAGAGAGTGCTCTTTGGAACAGAA	703
QY	166	IleValLeuProCysArgProProGluGlyIleProProAlaGluValGluTyrLeuArg	185
DB	704	GTCTTACTCCAGTGTGCACACCTGAAGGATCCAGTGGCTGAGGTGGAATGTTGAAA	763
QY	186	AsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSer	205
DB	764	AATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTATATTACTATTGTATCAAC	823
QY	206	LeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsn	225
DB	824	CTCATATAAGCAGGCGCCGACCTCTCTGATACTGCAAAATTACACCTGTGTGCCAAAC	883
QY	226	IleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTip	245
DB	884	ATTGTTGCCAAGAGAAAAGTAGCACTGCCACTGTCTATGTCTATGTCAACGGTGGCTGG	943
QY	246	SerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArg	265
DB	944	TCCACCTGGACGAGTGTCTGTGTGTACACGCGCTGTGGACGAGGTATCAGAAACGT	1003
QY	266	SerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGlnAsn	285
DB	1004	ACAAGACTTGTATCCAAACCGGACCACTCAATGGGGGTGCCCTTCTGTGAAGGCGAGT	1063
QY	286	ValHisArgThrValSerSerLeuLeuValSerValAspGlySerTrpSerProTip	305
DB	1064	GTG---CAGNAATAGCTGTACTACTATATGCCCAGTGGTGGCAGTGGACGCCATGG	1120
QY	306	SerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCysSerAsp	325
DB	1121	AGCAAGTGTGTCTATTGTGAACTGATGTCACCCACTTGGCGCAGGAGGAGTGACGCGC	1180
QY	326	ProAlaProArgAsnGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCys	345
DB	1181	CCAGCCCCCAAGATGGAGCAAGGACTGCGACGGCCCTCGTCTTGTCAATCCAGAACTGC	1240
QY	346	ThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyrValGly	365
DB	1241	ACTGATGGGCTTTGCGATGCGACTGCTCCTGATTGATGATGCTCTCTATGTTGGG	1300
QY	366	Leu---IleAlaValAlaValCysLeuValLeuLeuLeuValLeuLeuValLeuValTyr	384
DB	1301	ATTGTGATGACAGTGTGCTGCTGGCGATCTCTGTAGTTGTGGCCTGTTGTGTAT	1360
QY	385	CysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSerGly	404
DB	1361	CGAAGAATCATCGTCACTTTGATGAGATATTATTGACTCTTCGGCACTCAATGGGGC	1420
QY	405	PheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeuThrIleGln	424
DB	1421	TTTCACCTGTGAACATCAAGGAGCAAGACAGAT-----CTGCTGGCTGTATCCC	1471
QY	425	ProAspLeu---SerThrThrThrThrTyrGlnGlySerLeuCysProArgGlnAspGly	443
DB	1472	CCAGACTCAGTCAGCTGCAGCCATGTACAGAGGACCTGTCTATGCCCTGCATGAC---	1528
QY	444	ProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyArg	463
DB	1529	GTCTCAGACAAATCCCAATGACCACTCTCCAATTCTGGATCCACTGCCCAACCTGAAA	1588
QY	464	HisThrLeuHisHisSer-----SerProThrSerGluAlaGluGluPheVal	479
DB	1589	ATCAAGTGTACAAACCTCAGCTGCTGCTCCCCCAGATGACCTCTCTGAGTTACG	1648
QY	480	SerArgLeuSer-----ThrGlnAsnTyrPhe-----	488
DB	1649	TCCAAGCTGCTCCCTCAGATGACCCAGTCGTTCTGGAGAAATGAAGCCCTCAGCCTGAAG	1708
QY	489	---ArgSerLeuProArgGlyThr-----SerAsnMetThrTyrGlyThrPheAsn	505
DB	1709	AACGAGTGTACAAAGGAGACATGATCCATCTGACCCGATTTGGCAGCTTCAACTCG	1768
QY	506	LeuGlyClyArgLeuMetIleProAsnThrGlyIleSerLeuLeuLeuProAspAla	525
DB	1769	CTGGAGGTCACTTATTTGTTCCCAATTCAAGAGTCACTTGTGTATTCCTCGTGGGCGC	1828
QY	526	IleProArgLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArg	545
DB	1829	ATTCGCCAAGGAGAGTCTACGAAATGTATGTCTACTGTACACAGGAAGAACTATCAGG	1888
QY	546	LeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGly	565
DB	1889	CCACCATGGATGACTCTCAGACACTTTTGACCCCTGTGTGAGCTGTGGGCCCGCCAGG	1948
QY	566	ValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAsp	585
DB	1949	GCTCTGCTACCCGCGCCGCTGCTCTCACTATGCATCCTGCGCAGACCCCAATACCGAG	2008
QY	586	SerTrpSerLeuArgLeuLysLysGlnSerCysGlySerTrpGluGlnAspValLeu	605
DB	2009	GACTGAAATAATCTGCTCAAGAACCCAGGACAGCAGGACAGTGGAG---GATGTGGTG	2065
QY	606	HisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCys	625
DB	2066	GTGGTGGGGAGAGAAACTTCAACCCCTGCTGTACATTAAAGCTGATGTCAGAGGCTGC	2125
QY	626	TyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAla	645
DB	2126	CACATCTCACAGAGAACCTCAGCACTAGCCCTGGTAGGACATTCACCCACCACCAAGCG	2185
QY	646	AlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyr	665
DB	2186	GCTGCAAAAGCGCTCAAGCTGGCCATCTTTGGGCCCTCTGTGCTGCTCTCGCTGGAGTAC	2245
QY	666	AsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeu	685
DB	2246	AGCATCGAGTCTACTGTCTGGATGACCCAGGATGCCCTGAGGAATTTTACATCTT	2305
QY	686	GluLysGlnLeuGlyGlnLeuLeuGlnProArgValLeuHisPheLysAspSer	705

Db 2306 GAGAGACAGCGGAGGACAGCTCTAGAACCTTAAGGCTCTTCATTTTAAAGGCAGC 2365
 Qy 706 TyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeu 725
 Db 2366 ACCCACAACTGGCCCTGTCAATTCACGATATCCCATTCCTCTGGAAGAGCAATTC 2425
 Qy 726 LeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeu 745
 Db 2426 CTGGCTAAATATCAGGAATTCATTTTACCATGTTTGGAGTGGATCTCAAGAACAACCTG 2485
 Qy 746 HisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeu 765
 Db 2486 CACTGCACCTTCACTCTGGAAAGATTAGCTGTGAACACAGTGGAGCTGTTTGCAAACTC 2545
 Qy 766 TrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLys 785
 Db 2546 TGTGTGGCGGAGGTGGGAAGGAGGAGGAGGAGTCTTCCAGCTCAACTGCACCGTGTACAG 2605
 Qy 786 AspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValProAlaLeuVal 805
 Db 2606 GAACCTACTGGCATGATTGGCCGCTCTGGATCTCGACACCATCACACGGTCAAG 2665
 Qy 806 GlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleLeSerSerLeuAsp 825
 Db 2666 GGGCCCAAGTGTTCAGCATCCCTCTCCCTATCCGCGAGAGAGTCTCTAGCAGCTCGAT 2725
 Qy 826 ProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSer 845
 Db 2726 GCCCCCCAGACGAGGAGCCATGATGAGGATGTGGCCCATAGCTGAACCTGGACAGG 2785
 Qy 846 HisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGlu 865
 Db 2786 TACTGTAATTACTTGGCCCAATCAGGCCCACTGGGTAATCTCTGGATCTTTGGGAA 2845
 Qy 866 AlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGly 885
 Db 2846 GCACAGAACTTCCAGATGGAACCTGAGCATGTGGCAGTGTCTTGGGAAGAAATGGGA 2905
 Qy 886 GlnProAspAlaGlyLeuPheThrValSerGluAlaGlu 898
 Db 2906 AGACATGAACCGTGGTGTCTTAGCAGCAGAGGGGCGAG 2944
 RESULT 10
 ID ABK92062
 XX ABK92062 standard; DNA; 2995 BP.
 AC ABK92062;
 XX
 XX 14-AUG-2002 (first entry)
 XX
 XX DNA encoding novel UNC5 receptor-like protein.
 KW Cancer; metabolic disorder; skin disorder; infectious disease; anorexia;
 KW behavioral disorder; valve disease; endocrine disorder; inflammatory disorder;
 KW blood disorder; anxiety disorder; brain disorder; inflammatory disorder;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; dyslipidaemia;
 KW metabolic disturbance; obesity; metabolic syndrome X; wasting disorder;
 KW gene therapy; transgenic animal; human; gene; ds.
 XX
 XX Homo sapiens.
 OS
 PN W0200229058-A2.
 XX
 XX 11-APR-2002.
 PD
 XX
 XX 05-OCT-2001; 2001WO-US031248.
 XX
 XX 05-OCT-2000; 2000US-0238323P.
 PR 05-OCT-2000; 2000US-0238325P.
 PR 06-OCT-2000; 2000US-0238372P.
 PR 06-OCT-2000; 2000US-0238373P.

PR 06-OCT-2000; 2000US-0238379P.
 PR 06-OCT-2000; 2000US-0238382P.
 PR 06-OCT-2000; 2000US-0238383P.
 PR 06-OCT-2000; 2000US-0238384P.
 PR 06-OCT-2000; 2000US-0238397P.
 PR 06-OCT-2000; 2000US-0238400P.
 PR 06-OCT-2000; 2000US-0238401P.
 PR 06-OCT-2000; 2000US-0238402P.
 PR 14-MAR-2001; 2001US-0275892P.
 PR 08-JUN-2001; 2001US-0296860P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shimkets RA, Taupier RJ, Burgess CE, Zerhusen BD, Mezes PS;
 PI Rastelli L, Malyankar UM, Grosse WM, Alsobrook JP, Lepley DM;
 PI Spytek KA, Li L, Edinger S, Gerlach V, Ellerman K, Macdougall J;
 PI Gunther E, Millet I, Stone D, Smithson G, Szekeres ES;
 XX
 WPI: 2002-444103/47.
 DR P-PSDB; ABG61795.
 XX
 PT Novel isolated polypeptide, designated NOVX, useful for treating or
 PT preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and
 PT metabolic, neurodegenerative, immune and hematopoietic disorders.
 XX
 PS Claim 9; Page 93; 316pp; English.
 XX
 CC The invention describes an isolated polypeptide (I), useful in the
 CC manufacture of a medicament for treating a syndrome associated with a
 CC human disease. (I), the polynucleotide encoding it (II) and an antibody
 CC (III) to (I) are useful for treating or preventing cancer, metabolic
 CC disorders, skin disorders, infectious disease, anorexia, behavioral
 CC disorders, valve diseases, endocrine disorders, heart and blood
 CC disorders, anxiety disorders, brain disorders, inflammatory disorders,
 CC neurodegenerative disorder, Alzheimer's disease, Parkinson's disorder,
 CC immune disorders, haematopoietic disorders, and the various
 CC dyslipidaemias, metabolic disturbances associated with obesity, and the
 CC metabolic syndrome X and wasting disorders associated with chronic
 CC diseases. (I), (II) or (III) are useful in screening assays, detection
 CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
 CC predictive medicine (e.g., diagnostic assays and prognostic assays), and
 CC in methods of treatment (e.g., therapeutic and prophylactic). (I) is
 CC useful as immunogen to produce antibodies immunospecific for (I), to
 CC screen for potential agonist and antagonist compounds, and as bait
 CC protein in a two-hybrid or three-hybrid assay. (II) is useful in gene
 CC therapy, to express (I), to detect (I) mRNA or a genetic lesion in a (I)
 CC gene, and to modulate activity of (I). A cell containing a vector
 CC expressing (I) is useful for producing non-human transgenic animals. This
 CC sequence encodes a novel human polypeptide described in the invention
 XX
 SQ Sequence 2995 BP; 571 A; 984 C; 925 G; 515 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3,38e-162 Length: 2995
 Score: 2490.50 Matches: 496
 Percent Similarity: 67.86% Conservative: 150
 Best Local Similarity: 57.10% Mismatches: 247
 Query Match: 59
 DB: 6
 Gaps: 18
 US-09-970-944-2 (1-899) x ABK92062 (1-2995)
 QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaTrp--- 19
 Db 64 ATGGGGGCCCGAGCGAGCTCGGGCGCGCTGCTGCGACTGCTGCTCTCTCGGAC 123
 QY 20 -----LeuArgGlySerGlyAlaGlnGlnSerAlaThrVal---AlaAsnProValPro 36
 Db 124 CGAGGCTGAGCCAGCAGCAGGTAGGAGCGGTGAGGTGCTGCTGCTCTCTCCG 183
 QY 37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTrileVal 56
 Db 184 TCAGCGCCAGCAGCGCGCTGGCCTACTTCTTCGAGGAGCCACGAGCCGCTACATTGTG 243

QY 57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
DB 244 AAGAACAGCCTGTGGAGCTCCGTCGCGGCTTCCCGCCACACAGATCTACTTCAAG 303
QY 77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
DB 304 TGCNACGGCAGTGGGTACGCCAGAACGACCACTCACAGNAGGCGCTGGATGAGGCC 363
QY 97 Ser-----GlyGluProThrMetGluValArgIleAsnValSerArgGln 111
DB 364 ACCCTGGGGCGCGGGCGGCTCGCGGTGCGGAGGTGCGAGATCGAGGTGTCGCGCAG 423
QY 112 GlnValGlnLysValPheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSer 131
DB 424 CAGTGGAGGAGCTCTTTGGCTGGAGATTTACTGGTGGCAGTGGTGGCTGGAGCTCC 483
QY 132 SerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPhe 151
DB 484 CGGGCACCAACAGAGTGGCGAGCGCTAGCTCCGATCGCTCTGCGCAAGAACTTC 543
QY 152 GluGlnGluProLeuAlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArg 171
DB 544 GATCAGGAGCCTCTGGGCAAGAGGTGCCCTGGACCATGAGGTTCCTCTGCATGGCGC 603
QY 172 ProProGluGlyIleProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAsp 191
DB 604 CGCGCGAGGGGGCTGTGGCGGAGGTGAATGGCTCAAGAATGAGGATCATCGAC 663
QY 192 ProSerLeuAspProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAla 211
DB 664 CCCACCCAGGACACCACTTCTCTGCTCACCATCGACCAACCTCATCTCCGCGAGGCC 723
QY 212 ArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArg 231
DB 724 CGCTGTGCGACATGCGCAACTATACCTGCGTGGCGCAAGAACATCGTGGCCAAACGCGG 783
QY 232 SerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrp 251
DB 784 AGACCACTGCCACCGTCATCTGTCTAGTGAATGGCGGTGTCTCAGTGGCGAGATGG 843
QY 252 SerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsn 271
DB 844 TCACCTGTCTCCAAACCGCTGTGGCGAGGCTGGCAGAGCGCACCCGACCTGCACCAAC 903
QY 272 ProAlaProLeuAsnGlyIleAlaPheCysGluGlnAsnValHisAspArgThrVal 291
DB 904 CCCGCTCCACTCAACGAGGGGCTCTGCGAGGGCCAG---GCATTCCAGAAACCGGCC 960
QY 292 SerSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCys 311
DB 961 TGCACCACTCTGCCAGTCGATGGGGGTGGACGAGTGGACCAAGTGTGAGCTGC 1020
QY 312 GlyLeuAspCysThrHisTrpArgSerArgGluCysSerAspProAlaProArgAsnGly 331
DB 1021 AGCACTCAGTGTGCCACTGGCGTAGCGCGAGTGCATGGCGCCCGCCACCCCAAGACGGA 1080
QY 332 GlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysVal 351
DB 1081 GGCGGTGACTGCACGGGAGCGCTCTGACTCTAAGAACTGCACAGATGGGCTGTGCATG 1140
QY 352 HisSer-----AlaSerGlyProGluAspValAlaLeuTyr 363
DB 1141 CAAAGTGAGCTGTCCCGCAGTGTCTGGAGGCTCTCAGG-----GATCGCGGCTGTAT 1194
QY 364 ValGlyLeu-----IleAlaValAlaValCysLeuValLeuLeuLeuValLeuLeu 382
DB 1195 GCGGGCTCGTGGTGGCCATCTTCGTGTGTGCAATCCTCATGGCGGTGGGGGTGGT 1254
QY 383 ValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSer---IleLeu 401
DB 1255 GTGTACCGCCGCAACTGCGGTGACTTTCGACACAGACATCACTGACTCATCTGCTGCGCCTG 1314

QY 402 ThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu 421
DB 1315 ACTGGTGTTCACCCCGCTCACTTAAACGCGCAGGCCAGTAAACCCGAGCTCCTA 1374
QY 422 -----ThrIleGlnProAspLeuSerThrThr-----ThrTyrGlnGlySerLeuCys 438
DB 1375 CACCCCTCTGTGCTCTGACCTGACAGCCAGCGCCGCGCATCTACCGCGGACCGGTAT 1434
QY 439 ProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerPro 458
DB 1435 GCCTTGAGAGCTCC---ACCGACAAATCCCATGACCACTCTCTCTGCTGGAGCCC 1491
QY 459 LeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThr----- 472
DB 1492 TTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGCACCAGGGCTCTGGGCCAGGCGTG 1551
QY 473 SerGluAlaGluPheValSerArgLeuSerThrGlnAsnTyr----- 487
DB 1552 GCAGATGGGCTGACCTGTGGGGGTCTTGGCCGCTGGCACATACCCCTAGCGATTTCCGCC 1611
QY 488 -----PheArgSer----- 490
DB 1612 CGGACACCCACTTCTGTCACCTGCGCAGCGCCAGCCTCGGTTCGCGAGCTCTTGGGC 1671
QY 491 LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeu 510
DB 1672 CTGCCCCGACCCAGGAGCGGTCTGCGGCACCTTTGGCTGCTGGGTGGGAGGCTC 1731
QY 511 MetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLys 530
DB 1732 AGCATCCCGCGCAGGTGTGAGTGTGCTGCTGCCAATGAGGCATTTCCACAGGGCAAG 1791
QY 531 IleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla--- 549
DB 1792 TTCTACGAGATGATCTACTCATCAACAGCGAGAAAGTACC---CTGCCGCTTTCAGAA 1848
QY 550 GlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyValLeuLeuThr 569
DB 1849 GGGACCCAGACAGTATTGAGCCCTCGTGTGACCTGTGAGCCACAGGCGCTCTGCTGTC 1908
QY 570 ArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeu 589
DB 1909 CGCCCGCTCATCTCATCATGCCCCACTGTGCGAAGTCAGTGCCTGCTGATCTTT 1968
QY 590 ArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGlu 609
DB 1969 CAGCTCAAGACCCAGGCGCCACAGGGCCACTGGAGCAGAGGTGGTGCACCTGGATGAG 2028
QY 610 GluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThr 629
DB 2029 GAGACCCGTGAACACACCCCTGCTACTGCGAGCTGGAGCCCGAGGCTGTCACTCTGCTG 2088
QY 630 GluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaLysArg 649
DB 2089 GACCAGCTGGCGCCTACTAGTGTTCACGGGCGAGTCTTATTTCCCGCTCAGCAGTCAAGCGG 2148
QY 650 LeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgVal 669
DB 2149 CTCAGCTGGCGGCTCTTCGCCCCCGCCCTCTGCACTCCCTGGAGTACAGCTCCGGGTC 2208
QY 670 TyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeu 689
DB 2209 TACTGCTGGAGGACACGCTGTAGCACTCAAGGAGGTGTGGAGCTGGAGCGGACTCTG 2268
QY 690 GlyGlyGlnLeuIleGlnProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709
DB 2269 GCGGATATCTGTGGAGGAGCGCAACCGCTAATGTTCAAGGACAGTATCACCAACCTG 2328
QY 710 ArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyr 729
DB 2329 CGCTCTCCCTCCATGACCTCCCGCCATGCCCATTTGGAGGAGCAAGCTGCTGGCCAAATAC 2388
QY 730 GlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPhe 749

Db 2389 CAGGAGATCCCTCTCATCATTTGGAGTGGAGCCAGAGGCGCTCCATCTGACCTTC 2448
 Qy 750 ThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGln 769
 Db 2449 ACCCTGGAGAGGACAGCTTGGCCCTCCACAGAGCTCACTGCAAGATCTGGGTGGGCAA 2508
 Qy 770 ValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPhe 789
 Db 2509 GTGGAGGGAGGAGGCCAGATATTCAGCTGCATACCACTCTGGCA--GAGACACCTGCT 2565
 Qy 790 AlaGluLeuLeuAlaLeuGluSerGluAlaGly-----ValProAlaLeuValGlyPro 807
 Db 2566 GGCTCCCTGGACACTCTCTGCTCTGCGCCCTGGCAGCACTGTCAACCCAGCTGGGACCT 2625
 Qy 808 SerAlaPheLysIleProPheLeuIleArgGlnLysIleSerSerLeuAspProPro 827
 Db 2626 TATGCCCTTCAAGATPCCCACTGTCCATCCGCCAGAGATATGCAACAGCTAGATGCCCCC 2685
 Qy 828 CysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeu 847
 Db 2686 AACTCAGGGGCAATGACTGGCGATGTACACAGAAAGCTCTCTATGGACCGGTACTCTG 2745
 Qy 848 SerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArg 867
 Db 2746 AATTACTTTGCCACCAAGCGAGCCCGAGGGTGTGATCTCTGGACCTCTGGAGACTCTG 2805
 Qy 868 HisPheProAsnGlyAsnLeuSerGlnIleAlaAlaValAlaGlyLeuGlyGlnPro 887
 Db 2806 CACGAGACATGGGACCTCAACAGACCTGGCGAGTGCCTTTGGAGGAGATGGGCAAGAGT 2865
 Qy 888 AspAlaGlyLeuPheThrValSerGluAlaGluCys 899
 Db 2866 GAGATGCTGTGGCTGTGGCCACCGAGCGGAGCTGC 2901
 RESULT 11
 ABT06279
 ID ABT06279 standard; cDNA; 2860 BP.
 XX
 AC ABT06279;
 XX
 DT 24-OCT-2002 (first entry)
 XX
 DE Human NOV1a coding sequence.
 XX
 KW Human; autoimmune disease; cancer; infection; inflammatory disease;
 KW storage disorder; muscle disorder; neurodegenerative disorder; nontropic;
 KW developmental defect; neuroprotective; antiparkinsonian; hypotensive;
 KW hypertensive; haemostatic; cardiant; antianigmal; dermatological;
 KW immunosuppressive; antiinflammatory; virucide; antibacterial; anti-HIV;
 KW antiparasitic; antiallergic; antiasthmatic; antirheumatic; antiarthritic;
 KW vulnary; anorectic; antidiabetic; immunomodulator; antipsoriatic;
 KW nephrotropic; kerolytic; antitumor; cerebroprotective; anticonvulsant;
 KW antinfertility; antimanic; antidepressant; metabolic; cytotstatic;
 KW tranquilizer; analgesic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 DN WO200257450-A2.
 XX
 PD 25-JUL-2002.
 XX
 PF 29-NOV-2001; 2001WO-US048922.
 XX
 PR 29-NOV-2001; 2000US-0253834P.
 PR 30-NOV-2001; 2000US-0250926P.
 PR 25-JAN-2001; 2001US-0264180P.
 PR 20-AUG-2001; 2001US-0313656P.
 PR 05-OCT-2001; 2001US-0327456P.
 PR 28-NOV-2001; 2001US-00327456.
 XX
 (CURA-) CURAGEN CORP.

PI Edinger S, Macdougall JR, Millet L, Ellerman K, Stone DJ;
 PI Gerlach V, Grosse WM, Alsobrook JP, Lepley DM, Rieger D, Burgess CE;
 PI Caaman SJ, Spytek KA, Boldog FL, Li L, Padigaru M, Mishra V;
 PI Patturajan M, Shenoy S, Rastelli L, Tchernev VT, Vernet CAM;
 XX Zerhusen BD, Malyankar UM, Guo X, Miller CE, Gangolli EA;
 DR WPI: 2002-590741/63.
 DR P-FSDB: AAO18734.
 XX
 PT Novel isolated polypeptide, designated NOVX, useful for treating or
 PT preventing in NOVX-associated disorders e.g. cardiomyopathy,
 PT atherosclerosis, diabetes, cancer, allergy, asthma, Crohn's disease.
 XX
 PS Claim 9; Page 11; 353pp; English.
 XX
 CC The present invention provides the protein and coding sequences of
 CC several novel human proteins, designated NOVX. These can be used in the
 CC treatment of, amongst others, cancers, autoimmune diseases, infections,
 CC inflammatory diseases, storage disorders, muscle disorders,
 CC neurodegenerative diseases and developmental defects. The present
 CC sequence is a coding sequence of the invention
 XX
 SQ Sequence 2860 BP; 555 A; 950 C; 860 G; 495 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 8,93e-162 Length: 2860
 Score: 2494.00 Matches: 493
 Percent Similarity: 68.30% Conservative: 149
 Best Local Similarity: 58.45% Mismatches: 250
 Query Match: 51.89% Indels: 48
 DB: 6 Gaps: 18
 US-09-970-944-2 (1-899) x ABT06279 (1-2860)
 Qy 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrp--- 19
 Db 59 ATGGGGCCCGAGAGCGAGCTCGGGCGCGCTGCTGCTGCACTGTCTGCTGGGAC 118
 Qy 20 -----LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAlaProValPro 36
 Db 119 CCGAGGCTGAGCCAAAGCAGGCACTGATTTCTGGCAGCGAGGTGCTCCCTGACTCTCTCCG 178
 Qy 37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTrpIleVal 56
 Db 179 TCAGCGCCAGCAGAGCGCTGCGCTTCTCTCAGGAGCCACAGGACGCTACATTGTG 238
 Qy 57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
 Db 239 AAGAACAAGCTGTGGAGCTCCGCTCGCGCGCTTCCCGCCACACAGATCTACTTCAAG 298
 Qy 77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
 Db 299 TGAACGGCGAGTGGGTGACCAAGAACGACCACTCACACAGGAAGCCCTGGATGAGGCC 358
 Qy 97 SerGlyGluProThrMetGluValAlaGlyIleAsnValSerArgGlnValGluLysVal 116
 Db 359 ACCGGTCTCGGGGTGCGGAGTGCAGATCGAGGTGTCGCGCAGCAGGTGGAGGAGCTC 418
 Qy 117 PheGlyLeuGluGluTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLys 136
 Db 419 TTGGGCTGGAGGATTACTGGTGCAGTGGTGGCTTGGAGCTCCGGGGCACACCAAG 478
 Qy 137 SerGlnLysAlaTrpIleArgIleAlaArgLeuArgLysAsnPheGlnGluProLeu 156
 Db 479 AGTCGCGAGCGCTACGTCGCGATCGCTACCTCGCAAGAACTTCGATCAGGAGCCTCTG 538
 Qy 157 AlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProGluGlyIle 176
 Db 539 GGCAGGAGGTGCCCTGGACCATGAGGTTCCTCTGAGTGGCGCGCGGAGGGGTG 598
 Qy 177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196
 Db 599 CCTGTGGCGAGGTGGAAATGGCTCAAGAAATGAGGATGTATCATCGACCCACCCAGGACACC 658

Db 2798 GCCTTGAGGAGATGGGCACAGGTGAGATGCTGGTGGCTGTGGCCACGCGGACTGC 2857
 ||||| :
 RESULT 12
 ABT06280
 ID ABT06280 standard; cDNA; 2860 BP.
 XX
 AC ABT06280;
 XX
 DT 24-OCT-2002 (first entry)
 XX
 DE Human NOV1b coding sequence.
 XX
 KW Human; NOXV; autoimmune disease; cancer; infection; inflammatory disease;
 KW storage disorder; muscle disorder; neurodegenerative disorder; nontropic;
 KW developmental defect; neuroprotective; antiparkinsonian; hypotensive;
 KW hypertensive; haemostatic; cardiant; antiangiinal; dermatological;
 KW immunosuppressive; antiinflammatory; virucide; antibacterial; anti-HIV;
 KW antiparasitic; antiallergic; antiasthmatic; antineumatic; antiarthritic;
 KW vulnary; anorectic; antidiabetic; immunomodulator; antipsoriatic;
 KW nephrotrophic; kerolytic; antiulcer; cerebroprotective; anticonvulsant;
 KW antinfertility; antimanic; antidepressant; metabolic; cytostatic;
 KW tranquilizer; analgesic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 EN WO200257450-A2.
 XX
 PD 25-JUL-2002.
 XX
 PF 29-NOV-2001; 2001WO-US048922.
 XX
 PR 29-NOV-2000; 2000US-0253834P.
 PR 30-NOV-2000; 2000US-0250926P.
 PR 25-JAN-2001; 2001US-0264180P.
 PR 20-AUG-2001; 2001US-0313656P.
 PR 05-OCT-2001; 2001US-0327456P.
 PR 28-NOV-2001; 2001US-00327456.
 XX
 (CURA-) CURAGEN CORP.
 XX
 PA Edinger S, Macdonall JR, Millet I, Ellerman K, Stone DJ;
 PI Gerlach V, Grosse WM, Alsobrook JP, Lepley DM, Rieger D, Burgess CE;
 PI Casman SJ, Spytek KA, Boldog Fu, Li L, Padigaru M, Mishra V;
 PI Patturajan M, Shenoy S, Rastelli L, Tchernov VT, Vernet CM;
 PI Zerhusen BD, Malyankar UM, Guo X, Miller CE, Gangoli EA;
 XX
 DR WPI: 2002-590741/63.
 DR P-PSDB; AAO18735.
 XX
 PT Novel isolated polypeptide, designated NOXV, useful for treating or
 PT preventing in NOXV-associated disorders e.g. cardiomyopathy,
 PT atherosclerosis, diabetes, cancer, allergy, asthma, Crohn's disease.
 XX
 PS Claim 9; Page 14; 353pp; English.
 XX
 CC The present invention provides the protein and coding sequences of
 CC several novel human proteins, designated NOXV. These can be used in the
 CC treatment of, amongst others, cancers, autoimmune diseases, infections,
 CC inflammatory diseases, storage disorders, muscle disorders,
 CC neurodegenerative diseases and developmental defects. The present
 CC sequence is a coding sequence of the invention
 XX
 SQ Sequence 2860 BP; 557 A; 950 C; 858 G; 495 T; 0 U; 0 Other;

Alignment Scores:	
Pred. No.:	8.93e-162
Score:	2494.00
Percent Similarity:	68.30%
Best Local Similarity:	52.45%
Query Match:	51.89%
DB:	6
	Gaps:
	Indels:
	Mismatches:
	Conservative:
	Matches:
	Length:
	2860

QY 356 GlyProGluaspValAlaLeuTyrValGlyIleu---IleAlaValAlaValCysLeuVal 374
Db 1136 GGG-----GATGGGGCTGTATGCGGGGCTGTGGTGGCCATCTTCTGGTGGTGGCA 1189
QY 375 LeuLeuLeuValLeuLeuValTyrCysArgLysLysGluGlyLeuAspSerAsp 394
Db 1190 ATCTCTATGGCGGTGGGGTGGTGTACCGCGCAACTGCCGTGACTTCGACACAGAC 1249
QY 395 ValAlaaspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSer 413
Db 1250 ATCACTGACTCATCTGCTGCGCTGACTGGTGGTTCACCGCGTAACTTTAAGACGCA 1309
QY 414 LysAlaaspAsnProHisLeuLeu-----ThrIleGlnProAspLeuSerThrThr 431
Db 1310 AGCCCCAGTAACCGCGAGCTCTACACCCCTCTGTGCTCTGTGACTGACGACGAGCGCC 1369
QY 432 ---ThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeu 450
Db 1370 GGCATCTACCGCGGACCGGTGTATGCCCTGCAGGACTCC---ACCGACAAATCCCATG 1426
QY 451 ThrAsnGlyHisLeuLeuSerProLeuGlyGlyArgHisThrLeuHisHisSerSer 470
Db 1427 ACCAACTCTCTGCTGGACCCCTTACCAGCCCTTAAGGTCAAGTCTACAGCTCCAGC 1486
QY 471 ProThr-----SerGluAlaGluGluPheValSerArgLeuSerThr 484
Db 1487 ACCACGGGCTGTGGCCAGGCTGGCAGATGGGGCTGACTGCTGGGGTCTTCCCGCT 1546
QY 485 GlnAsnTyr-----LeuProArgGlyThrSerAsnMetThrTyrGlyThr 502
Db 1547 GGCACATACCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGGCGAGCGCCAGC 1606
QY 491 -----LeuProArgGlyThrSerAsnMetThrTyrGlyThr 502
Db 1607 CTCGGTTCGCCAGGACTCTTGGGGCTGCCCGAGACCGGAGCAGCGTTCAGCGGCA 1666
QY 503 PheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePro 522
Db 1667 TTTGGCTGCTGGTGGGAGGCTCAGCATCCCGCGCAGGGGGTCACTGCTGGTGGCC 1726
QY 523 ProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGlu 542
Db 1727 AATGGAGCCATTCCCGAGGCAAGTTCTACAGATGTATCTACTCATCAACAGGCGAA 1786
QY 543 AspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSerProIleValSerCys 561
Db 1787 AGTACC---CTGCCGCTTTCAGAAAGGACCCAGACAGTATTGAGCCCTCGTGACCTGT 1843
QY 562 GlyProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGlu 581
Db 1844 GGACCCACAGGCTCTGCTGTGGCGGCGGTATCTCTACCATGCCCTGCTGGCGAA 1903
QY 582 ProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGlu 601
Db 1904 GTCACTGCCGTGACTGATCTTTCAGCTCAAGACCCAGGCGCCACAGGGCCACTGGGAG 1963
QY 602 GlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeuGlu 621
Db 1964 ---GAGGTGGTGAACCTGGATGAGGAGACCCCTGAACACACCCCTGCTACTGCGCAGCTGGAG 2020
QY 622 AlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAla 641
Db 2021 CCAGGGCTGTACATCTCTGCTGGACGAGTGGGACCTACGTGTTCCAGGGGAGTCC 2080
QY 642 LeuSerValAlaAlaLysArgLeuLysLeuLeuPheAlaProValAlaCysThr 661
Db 2081 TATTCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCGGTCTTCGCCCGCCCTCTGCACC 2140
QY 662 SerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGlu 681
Db 2141 TCCCTGGAGTACACCTCCGGGTCTACTGCTGGAGGACACGCTGTAGCACTGAAGGAG 2200

QY 682 ValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHis 701
Db 2201 GTCTGAGCTGGAGCGGACTCTGGCGGATACTTGGTGGAGGAGCCGAACCCCTATG 2260
QY 702 PheLysaspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrp 721
Db 2261 TTCAAGGACAGTTACCACAACTTGGCGCTCTCCCTCCATGACCTCCCCCAATGG 2320
QY 722 LysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThr 741
Db 2321 AGAGGCAAGTCTGGCCAAATACAGAGATCCCTTCTATCACATTTGGAGTGGCAGC 2380
QY 742 GlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeu 761
Db 2381 CAGAAGGCCCTCCACTGCATTTTCACTCCCTGGAGGACAGCTTGGCTCCACAGAGCTC 2440
QY 762 AlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPhe 781
Db 2441 ACCTGCAAGATCTCGCTGGCAAGTGAAGGGGAGGCGCAGATATTCAGCTGCATACC 2500
QY 782 AsnIleThrLysaspThrArgPheAlaGluLeuAlaLeuGluSerGluAlaGly--- 800
Db 2501 ACTCTGGCA---GAGACACCTGTGGCTCCCTGGACACTCTCTGCTCTGCCCTGGCAGC 2557
QY 801 ---ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db 2558 ACTGTCAACCCAGCTGGGACCTTATGCCCTTCAAGATCCCACTGTCTCATCCGCCAGAG 2617
QY 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db 2618 ATATGCAACAGCTTAGATGCCCACTCAACGCGGCAATCACTGGCGGATTTAGCAGAG 2677
QY 840 LysLeuHisLeuaspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db 2678 AAGCTCTTATGGACCGGTACTGTAATTCTTTCACCAAGCGAGCCCAACCGGTGTG 2737
QY 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db 2738 ATCTTGACCTCTGGGAGGCTCTGCAGCAGGACGATGGGACCTCAACAGCTTGGCGAGT 2797
QY 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 2798 GCCTTGGAGGAGATGGCAAGAGTGAGATGTGTGTGGCTGTGGCCACCGCGGAGTGC 2857
RESULT 13
AAS21316
ID AAS21316 standard; cDNA; 3884 BP.
XX AC AAS21316;
XX DT 24-OCT-2001 (first entry)
XX DE Human cDNA sequence encoding for PRO4326 polypeptide.
XX KW Human secretory and transmembrane; PRO; mammalian; cancer; lung; breast;
KW prostate; cervical; tumour necrosis factor-alpha; TNF-alpha; cartilage;
KW ear; proliferation; glucose; free fatty acid; skeletal muscle; adipocyte;
XX A-peptide; factor VIIA; gene therapy; ss.
XX OS Homo sapiens.
XX FN WO200140466-A2.
XX PD 07-JUN-2001.
XX 01-DEC-2000; 2000WO-US032678.
XX PF 01-DEC-1999; 99WO-US028301.
XX PR 01-DEC-1999; 99WO-US028634.
XX PR 02-DEC-1999; 99WO-US028551.
XX PR 02-DEC-1999; 99WO-US028564.
XX PR 02-DEC-1999; 99WO-US028565.
XX PR 09-DEC-1999; 99US-0170262P.

PR 16-DEC-1999; 99WO-US0300095.
 PR 20-DEC-1999; 99WO-US030911.
 PR 20-DEC-1999; 99WO-US030919.
 PR 30-DEC-1999; 99WO-US031243.
 PR 30-DEC-1999; 99WO-US031274.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000277.
 PR 06-JAN-2000; 2000WO-US000376.
 PR 11-FEB-2000; 2000WO-US003565.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 18-FEB-2000; 2000WO-US004342.
 PR 22-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US004914.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 01-MAR-2000; 2000WO-US005601.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 03-MAR-2000; 2000US-0187202P.
 PR 10-MAR-2000; 2000WO-US006319.
 PR 15-MAR-2000; 2000WO-US006884.
 PR 20-MAR-2000; 2000WO-US007377.
 PR 21-MAR-2000; 2000WO-US007532.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 05-JUN-2000; 2000US-0209832P.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 11-AUG-2000; 2000WO-US022031.
 PR 23-AUG-2000; 2000WO-US023522.
 PR 24-AUG-2000; 2000WO-US023328.
 PR 08-NOV-2000; 2000WO-US030952.
 PR 10-NOV-2000; 2000WO-US030873.
 XX
 PA (GETH) GENENTECH INC.

XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 XX
 DR WPI: 2001-408281/43.
 DR P-PSDB; AAU12244.

XX Isolated, secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing PRO
 PT polypeptides, and detect the presence of mammalian tumors e.g. lung,
 PT breast, prostate, cervical.

XX Claim 3; Fig 145; 813pp; English.

XX AAS21244-AAS21518 encode for novel human secretory and transmembrane PRO
 CC polypeptides. The PRO polypeptides are useful to detect other PRO
 CC polypeptides, to link bioactive molecules to cells expressing PRO
 CC polypeptides, to modulate biological activities of cells expressing PRO
 CC polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumors by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample. Some
 CC of the 275 sequences are also useful to stimulate the release of tumour
 CC necrosis factor-alpha (TNF-alpha) from human blood, the proliferation or
 CC differentiation of chondrocytes, the proliferation or gene expression in
 CC pericyte cells, the release of proteoglycans from cartilage, the
 CC proliferation of inner ear utricular supporting cells or of T-
 CC lymphocytes, the release of a cytokine from peripheral blood monocytes
 CC (PBMCs), or the proliferation of endothelial cells. Some of the PRO
 CC polypeptides may modulate glucose or free fatty acid uptake by skeletal
 CC muscle cells or by adipocytes; or inhibit binding of A-peptide to factor
 CC VIIA. The PRO polypeptides can be used in assays to identify molecules
 CC involved in binding interactions. The polynucleotides encoding PRO
 CC polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy
 XX
 SQ Sequence 3884 BP; 767 A; 1278 C; 1162 G; 677 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2,96e-161 Length: 3884
 Score: 2479.00 Matches: 493
 Percent Similarity: 67.54% Conservative: 150
 Best Local Similarity: 51.79% Mismatches: 249
 Query Match: 51.79% Indels: 60
 DB: 4 Gaps: 18

US-09-970-944-2 (1-899) x AAS21316 (1-3884)

Qy 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAaTrp--- 19
 Db ATGGGGCCCGAGCGAGCTCGGGGCGCGTGTGTGTGGCACTGCTGTGTGGGAC 457
 Qy 20 -----LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAaProValPro 36
 Db CCGAGGCTGAGCAAGCAGGCACTGATTCTGGCAGCGAGGTGTCTCCCTGACTCTCTCCG 517
 Qy 37 GlyAlaAaAsnProAspLeuProHisPheLeuValGluProGluAspValTyrIleVal 56
 Db TCAGCGCCAGCAGCGCGCTGCTCTTCTTGAGGAGCCACAGGACGCTACATTGTG 577
 Qy 57 LysAsnLysProValLeuLeuValCysLeuAlaValProAlaThrGlnIlePhePheLys 76
 Db AAGAACAAGCCTGTGGAGCTCGCTGCGCGCCTTCCCGCCACACAGATCTACTTCAAG 637
 Qy 77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
 Db TGAACCGCGAGTGGGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 697
 Qy 97 SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysVal 116
 Db ACCGCGCTCGGGTGGCGAGGTGAGTGTGCGGAGGAGGTGAGGTGAGGTGAGGTGAG 757
 Qy 117 PheGlyLeuGluGluTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLys 136
 Db TTTGGGTGAGGATTAATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 817
 Qy 137 SerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeu 156
 Db AGTCGCGCGAGCTACGTCCGATCGCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTG 877
 Qy 157 AlaLysValSerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIle 176
 Db GGCAGAGGTGCGCTGGAGCATGAGGTCTCTGCGAGTGGCGCGCGCGAGGGGTG 937
 Qy 177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196
 Db CCTGTGCGCGAGGTGGAATGGCTCAAGATGAGGATGTCATCGACCCACCCAGGACCC 997
 Qy 197 AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216
 Db AACTTCTGCTACCATCGACCAACACCTCATCGCCAGGCGCGCGCTGTGGACACT 1057
 Qy 217 AlaAsnTrpThrCysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAla 236
 Db GCCACTATACCTGCTGGCCCAAGAACTGTGGCAAAACATGTGGCAAAACATGTGGCA 1117
 Qy 237 ValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAla 256
 Db GTCATGCTACGTGAATGGCGCTGGTTCAGCTGGGAGAGTGGTTCACCTGCTCAAC 1177
 Qy 257 SerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276
 Db CGTGTGGCGAGGCTGGCAGAAAGCGCACCGGACCTGACCAACCCCGCTGCCCACTCA 1237
 Qy 277 GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuVal 296
 Db GGAGGGGCGCTTTCGCGAGGGCCAG---GCATTCCAGAAAGACCCGCTGCACCACTTCG 1294
 Qy 297 SerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThr 316
 Db CCAGTGCATGGGCGGTGGAGGAGTGGCAACTGTGTGAGCTGAGCAGCTGAGTGTGCC 1354

Qy 317 HistProArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGln 336
Db 1355 CACTGGGTAGCCGCGAGTGCATGCGCCGCCACCCAGAACGAGCGCGTGAAGTGCAGC 1414
Qy 337 GlyThrAspLeuAspThrArgAsnGlySerAspLeuGlyCysValHisSer--- 353
Db 1415 GGGACGCTGCTCAGACTCTAAGACTGACAGATGGGCTGTGTCATGCAAAATAAGAAACT 1474
Qy 354 -----AlaSerGlyProGluAspValAlaLeuTyr 363
Db 1475 CTAAGCGACCCCAACAGCACCTGCTGGAGCCCTCAGG-----GATGCGGCTGTAT 1528
Qy 364 ValGlyLeu---IleAlaValAlaValCysLeuValLeuLeuLeuValLeuLeu 382
Db 1529 GCGGGGCTCGTGGGGCCATCTTCGTGTCGTCGAATCTCATGGCGGTGGGGTGTG 1588
Qy 383 ValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSer---IleLeu 401
Db 1589 GTGTACCGCGCAACTGCCGTGACTTCGACACACATCAGTACTATCTGCTGCCCTG 1648
Qy 402 ThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu 421
Db 1649 ACTGGTGGTTCCACCCCGTCAACTTAAAGACGGCAAGGCCAGCAACCCGACGCTCCTA 1708
Qy 422 -----ThrIleGlnProAspLeuSerThrThrThr-----ThrTyrGlnGlySerLeuCys 438
Db 1709 CACCCCTCTGTCCTCTCCAGCTCAGACCGCGCGGCACTACCGCGGACCCGCTGTAT 1768
Qy 439 ProArgGlnAspGlyProSerProLysPheGlnLeuThrAenGlyHisLeuLeuSerPro 458
Db 1769 GCCCTGCAGGACTCC---ACGACAAAATCCCAACCACTCTCCTCTGCTGGACCCC 1825
Qy 459 LeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThr----- 472
Db 1826 TTATCCAGCCCTTAAGGTCAAGTCTACAGCTCCAGCACACCGGCTCTGGGCCAGGCGCTG 1885
Qy 473 SerGluAlaGluGluPheValSerArgLeuSerThrGlnAenTyr----- 487
Db 1886 GCAGATGGGGTGACCTGCTGGGGTCTTGGCGCTGGCACATACCCCTAGGATTTGCCC 1945
Qy 488 -----PheArgSer----- 490
Db 1946 CGGGACACCCACTTCTCCTGACCTGGCAGCGCCAGCCTCGTTCCTCCAGCAGCTCTTGGGC 2005
Qy 491 LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeu 510
Db 2006 CTGCCCCGAGACCCAGGAGCAGCGTCAAGCGGCACCTTTGGCTGCTGGTGGGAGGCTC 2065
Qy 511 MetIleProAenThrGlyIleSerLeuLeuLeuProProAspAlaIleProArgGlyLys 530
Db 2066 AGCATCCCGGCACAGGGGTGAGTGTGTCGCCCAATGAGGACATTCCTCCAGGGCAAG 2125
Qy 531 IleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla--- 549
Db 2126 TTCTACGAGATGTATCTACTCATCAACAAGCAGAAAGTACC---CTCCCGCTTTCAGAA 2182
Qy 550 GlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyValLeuLeuThr 569
Db 2183 GGGACCCAGACAGATTGAGCCCTCGGTGACCTGTGGACCCACAGGCGCTCTGCTGTGC 2242
Qy 570 ArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTyrSerLeu 589
Db 2243 CGCCCGCTCATCTTCACATGCCACCTGTCGCCAAGTCACTGCGCGTGAAGTCTTT 2302
Qy 590 ArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGlu 609
Db 2303 CAGCTCAAGACCCAGGCCACAGGGCCACTGGGAG---GAGGTGGTGACCTCGATGAG 2359
Qy 610 GluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThr 629
Db 2360 GAGACCCCTGAACACACCCCTGCTACTGTCAGCTGAGGCCCGAGGGCTGTCACATCTGCTG 2419

Qy 630 GluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaLysArg 649
Db 2420 GACCAGCTGGCCACCTACGTGTTCCACGGCGAGTCTTATTCCTCCCTCAGCAGTCAAGCGG 2479
Qy 650 LeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgVal 669
Db 2480 CTCAGCTGGCGCTTTCGCCCGCCCTCTGCACCTCCCTGGAGTAGTACAGCTCCGGGTC 2539
Qy 670 TyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeu 689
Db 2540 TACTGCTGGAGGACCGCTGTAGCACCTGAAGAGGTGTGGAGCTGGAGCGGACTCTG 2599
Qy 690 GlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709
Db 2600 GCGGATACCTTGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACCAACCTG 2659
Qy 710 ArgLeuSerIleHisAspValProSerSerLeuTyrLysSerLysLeuValSerTyr 729
Db 2660 CGCCTCTCCCTCCATGACCTCCCATGTCATGGAGGAGCAAGCTGCTGGCCAAATAC 2719
Qy 730 GlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPhe 749
Db 2720 CAGGAGATCCCTTCTATCATATTGAGTGGCAGCAGAGGCCCTCCACTGACATTC 2779
Qy 750 ThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTyrValTyrGln 769
Db 2780 ACCGTGAGAGGCACAGCTTGGCTCCACAGAGCTCACCTGCAAGATCTGGTGGCGCAA 2839
Qy 770 ValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPhe 789
Db 2840 GTGAAGGGAGGGCCAGCATATTCAGCTGCATACCATCTCTGGCA---GAGACACCTGCT 2896
Qy 790 AlaGlnLeuLeuAlaLeuGluSerGluAlaGly-----ValProAlaLeuValGlyPro 807
Db 2897 GGCTCCCTGGACACTCTCTGCTCTGCCCCCTGGCAGCACTGTACACCCAGCTGGGACCT 2956
Qy 808 SerAlaPheLysIleProPheLeuIleArgGlnLysIleLeuSerSerLeuAspProPro 827
Db 2957 TATGCTTCAAGATCCCACTGCTCCATCGCCAGAGATATGCAACAGCTAGATGCCCCC 3016
Qy 828 CysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeu 847
Db 3017 AACTCACGGGCAATGACTGGCGATGTTAGCACAGAAGCTCTCTATGGACCGGTACCTG 3076
Qy 848 SerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTyrGluAlaArg 867
Db 3077 AATTACTTTGCCAACAAAGCGAGCCCAACCGGTGTGATCTCGGACCTCTGGAAAGCTGT 3136
Qy 868 HisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnPro 887
Db 3137 CAGCAGCAGATGGGACCTCACAGCTGGCGAGTGCCTTGGAGGAGATGGGCAAGAGT 3196
Qy 888 AspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 3197 GAGATGCTGGTGGTGTGGCCACCGAGCGGAGCTGC 3232
RESULT 14
ACD23925
ID ACD23925 standard; cDNA; 3884 BP.
XX
AC ACD23925;
XX
DT 26-AUG-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO4326 cDNA.
XX
KW Human; secreted and transmembrane protein; PRO; anti-inflammatory;
KW antiarteriosclerotic; cardiant; anti-infertility; anti-HIV; cytostatic;
KW antidiabetic; gene therapy; tumour necrosis factor (TNF)-alpha release;
KW TNF-alpha release; cell proliferation; cell differentiation;
KW gene expression modulator; proteoglycan release; cytokine release;
KW tumour; inflammatory disease; organ failure; atherosclerosis;
KW cardiac injury; infertility; birth defect; premature aging; AIDS;

acquired immunodeficiency syndrome; cancer; diabetic complication;
chromosome mapping; gene mapping; pharmaceutical; diagnostic; biosensor;
bioreactor; tissue typing; gene; ss.

Homo sapiens.

US2003032156-A1.

13-FEB-2003.

06-MAY-2002; 2002US-00140474.

31-MAR-1997; 97WO-US005230.

12-JUN-1998; 98WO-US012456.

14-JUL-1998; 98WO-US014552.

28-AUG-1998; 98WO-US017888.

10-SEP-1998; 98WO-US018824.

14-SEP-1998; 98WO-US019093.

14-SEP-1998; 98WO-US019094.

16-SEP-1998; 98WO-US019177.

17-SEP-1998; 98WO-US019330.

07-OCT-1998; 98WO-US021141.

29-OCT-1998; 98WO-US022991.

20-NOV-1998; 98WO-US022992.

01-DEC-1998; 98WO-US025108.

05-JAN-1999; 99WO-US000106.

08-MAR-1999; 99WO-US005028.

10-MAR-1999; 99WO-US005190.

20-APR-1999; 99WO-US008615.

14-MAY-1999; 99WO-US010733.

02-JUN-1999; 99WO-US012252.

01-SEP-1999; 99WO-US020111.

08-SEP-1999; 99WO-US020594.

13-SEP-1999; 99WO-US020944.

15-SEP-1999; 99WO-US021090.

05-OCT-1999; 99WO-US021547.

29-NOV-1999; 99WO-US023089.

30-NOV-1999; 99WO-US028214.

30-NOV-1999; 99WO-US028313.

01-DEC-1999; 99WO-US028409.

01-DEC-1999; 99WO-US028301.

02-DEC-1999; 99WO-US028634.

02-DEC-1999; 99WO-US028551.

02-DEC-1999; 99WO-US028564.

02-DEC-1999; 99WO-US028565.

16-DEC-1999; 99WO-US030095.

20-DEC-1999; 99WO-US030911.

20-DEC-1999; 99WO-US030999.

22-DEC-1999; 99WO-US030720.

30-DEC-1999; 99WO-US031243.

30-DEC-1999; 99WO-US031274.

05-JAN-2000; 2000WO-US000219.

06-JAN-2000; 2000WO-US000277.

06-JAN-2000; 2000WO-US000376.

11-FEB-2000; 2000WO-US003565.

18-FEB-2000; 2000WO-US004341.

18-FEB-2000; 2000WO-US004342.

22-FEB-2000; 2000WO-US004414.

24-FEB-2000; 2000WO-US004914.

24-FEB-2000; 2000WO-US005004.

01-MAR-2000; 2000WO-US005601.

02-MAR-2000; 2000WO-US005746.

02-MAR-2000; 2000WO-US005841.

10-MAR-2000; 2000WO-US006319.

15-MAR-2000; 2000WO-US006884.

20-MAR-2000; 2000WO-US007377.

21-MAR-2000; 2000WO-US007532.

30-MAR-2000; 2000WO-US008439.

17-MAY-2000; 2000WO-US013705.

22-MAY-2000; 2000WO-US014042.

PR 30-MAY-2000; 2000WO-US014941.

PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 01-MAR-2001; 2001WO-US006520.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 18-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 01-JUN-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 05-JUN-2001; 2001WO-US017800.
PR 14-JUN-2001; 2001US-00874503.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
XX
PA (GETH) GENENTECH INC.

Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;

Gerriksen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;

Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;

WPI; 2003-341980/32.

P-PSDB; ABO17688.

New secreted and transmembrane PRO nucleic acids, for treating

inflammation, organ failure, atherosclerosis, cardiac injury,

infertility, birth defects, premature aging, acquired immunodeficiency

syndrome (AIDS), or cancer.

Claim 2; Fig 145; 660pp; English.

The invention describes an isolated nucleic acid (I) comprising, or which

has 80 % sequence identity to, or the full-length coding sequence of, one

of 275 nucleotide sequences, and which encodes a corresponding

polypeptide selected from 275 amino acid sequences, where all sequences

are given in the specification. The polypeptide encoded by (I) is used to

detect PRO polypeptides, link a bioactive molecule to a cell expressing a

PRO polypeptide, modulate a biological activity of a cell, stimulate the

release of tumour necrosis factor (TNF)-alpha from human blood, modulate

the uptake of glucose or free fatty acid by cells, stimulate or inhibit

the proliferation or differentiation of cells or gene expression,

stimulate the release of proteoglycans, stimulate the release of cytokine

from peripheral blood mononuclear cells, inhibit the binding of A-peptide

acid and polypeptide encoded by it, are useful for treating inflammatory

diseases, organ failure, atherosclerosis, cardiac injury, infertility,

birth defects, premature aging, acquired immunodeficiency syndrome

(AIDS), cancer, or diabetic complications. The nucleic acid is useful as

hybridisation probes, in chromosome and gene mapping, and in generating


```
Db 2303 CAGCTCAAGACCCAGGCCCCCAGGCGCCACTGGAG--GAGTGTGACCTGGATGAG 2359
QY 610 GluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThr 629
Db 2360 GAGACCTGACACACCTGCTACTGCCAGCTGGAGCCAGGCGCTGCATCTCTGCTG 2419
QY 630 GluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaValArg 649
Db 2420 GACCAGCTGGACACTACGTGTTCCGGGGGAGTCTATTCCGCTCAGCATCAAGCGG 2479
QY 650 LeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgVal 669
Db 2480 CTCACGCTGGCGCTCTTCGCCCGCGCTCTGCACCTCCCTGGAGTACAGCCTCGGGTC 2539
QY 670 TyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeu 689
Db 2540 TACTGCTGGAGACACGCTGTAGCACTGAAGGAGTCTGGAGCTGGAGCGGACTCTG 2599
QY 690 GlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709
Db 2600 GCGGATCTTGGTGAGAGCGCGAAACCGCTTAATGTTCAAGACAGTTACACACCTG 2659
QY 710 ArgLeuSerIleHisAspValProSerSerLeuTyrLysSerLysLeuValSerTyr 729
Db 2660 CGCCTCTCCCTCCATGACCTCCCGCATGCCATTTGGAGGAGCAAGCTGTGGCCAAATAC 2719
QY 730 GlnGluIleProPheTyrHisIleTyrAsnGlyThrGlnArgTyrLeuHisCysThrPhe 749
Db 2720 CAGGAGATCCCTTCTATCACATTTGGAGTGAGCAGCCAGAGCCCTCCACATGCACTTC 2779
QY 750 ThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTyrValTrpGln 769
Db 2780 ACCCTGGAGAGCAGACTTGGCTCCAGAGCTCACTGCAAGATCTGCGTGGCGCAA 2839
QY 770 ValGluGlyAspGlyGlnSerPheSerIleAsnIleThrLysAspThrArgPhe 789
Db 2840 GTGGAAGGAGGCGGCAGATATTCAGCTGCATACCACCTCTGGCA--GAGACACCTGCT 2896
QY 790 AlaGluLeuLeuAlaLeuGluSerGluAlaGly-----ValProAlaLeuValGlyPro 807
Db 2897 GCTCCCTGGACACTCTCTGCTCTGCCCTGGCAGCACTGTACCACCCAGCTGGGACCT 2956
QY 808 SerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProPro 827
Db 2957 TATGCTTCAAGATCCCACTGCTCCATCCGCCAGAGATATGCAACAGCTAGATCCCCCC 3016
QY 828 CysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeu 847
Db 3017 AACTCAGGGGCAATGACTGGCGGATGTTAGCACAGAAGCTCTCTATGGACCGGTACCTG 3076
QY 848 SerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArg 867
Db 3077 AATTACTTGGCCACCAAGCGAGCCCGGCGGTGTGATCTCGGACTCTCTGGGAAGCTCTG 3136
QY 868 HisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnPro 887
Db 3137 CAGCAGGAGCATGGGACCTCAACAGCTGGCGAGTCTTGGAGGAGATGGGCAAGT 3196
QY 888 AspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 3197 GAGATGCTGGTGGCTGTGGCCACCGGAGGCGGACTGCG 3232
RESULT 15
ACA67066
ID ACA67066 standard; cDNA; 3884 BP.
XX
AC ACA67066;
XX
DT 23-JUN-2003 (first entry)
XX
DE cDNA encoding human PRO polypeptide #73.
XX
```

```
KW Human; PRO polypeptide; secreted and transmembrane protein;
KW anti-PRO antibody; diagnostic assay; gene expression; diabetes;
KW bone disorder; cartilage disorder; rheumatoid arthritis; obesity;
KW sports injury; osteoarthritis; hyper-insulinaemia; hypo-insulinaemia;
KW hearing loss; coagulation disorder; stroke; heart attack; cardiometabolic;
KW antidiabetic; anorectic; vulnervary; antiarthritic; osteopathic;
KW antirheumatic; auditory; cerebroprotective; angiogenic; gene; ss.
XX Homo sapiens.
OS
XX
PN US2003004311-A1.
XX
PD 02-JAN-2003.
XX
PF 19-DEC-2001; 2001US-00028072.
XX
PR 18-JUN-1997; 97US-0049911P.
PR 26-AUG-1997; 97US-0056974P.
PR 17-SEP-1997; 97US-0059113P.
PR 17-SEP-1997; 97US-0059115P.
PR 17-SEP-1997; 97US-0059117P.
PR 17-SEP-1997; 97US-0059122P.
PR 17-SEP-1997; 97US-0059184P.
PR 18-SEP-1997; 97US-0059263P.
PR 19-SEP-1997; 97US-0059352P.
PR 19-SEP-1997; 97US-0059588P.
PR 19-SEP-1997; 97US-0059836P.
PR 17-OCT-1997; 97US-0062250P.
PR 17-OCT-1997; 97US-0062285P.
PR 17-OCT-1997; 97US-0062287P.
PR 17-OCT-1997; 97US-0063755P.
PR 24-OCT-1997; 97US-0062814P.
PR 24-OCT-1997; 97US-0062816P.
PR 24-OCT-1997; 97US-0063045P.
PR 24-OCT-1997; 97US-0063082P.
PR 24-OCT-1997; 97US-0063127P.
PR 27-OCT-1997; 97US-0063327P.
PR 27-OCT-1997; 97US-0063329P.
PR 28-OCT-1997; 97US-0063550P.
PR 28-OCT-1997; 97US-0063561P.
PR 29-OCT-1997; 97US-0063704P.
PR 29-OCT-1997; 97US-0063733P.
PR 29-OCT-1997; 97US-0063735P.
PR 29-OCT-1997; 97US-0063738P.
PR 03-NOV-1997; 97US-0064248P.
PR 12-NOV-1997; 97US-0064809P.
PR 12-NOV-1997; 97US-0065186P.
PR 21-NOV-1997; 97US-0065846P.
PR 21-NOV-1997; 97US-0066364P.
PR 24-NOV-1997; 97US-0066453P.
PR 24-NOV-1997; 97US-0066511P.
PR 24-NOV-1997; 97US-0066770P.
PR 11-DEC-1997; 97US-0069212P.
PR 11-DEC-1997; 97US-0069278P.
PR 11-DEC-1997; 97US-0069334P.
PR 16-DEC-1997; 97US-0069694P.
PR 23-JAN-1998; 98US-0072320P.
PR 04-FEB-1998; 98US-0073612P.
PR 09-FEB-1998; 98US-0074086P.
PR 12-MAR-1998; 98US-0074092P.
PR 20-MAR-1998; 98US-0077791P.
PR 25-MAR-1998; 98US-0079294P.
PR 27-MAR-1998; 98US-0079663P.
PR 31-MAR-1998; 98US-0079728P.
PR 12-JUN-1998; 98US-0080165P.
PR 14-JUL-1998; 98WO-US012456.
PR 28-AUG-1998; 98WO-US014552.
PR 10-SEP-1998; 98WO-US017888.
PR 14-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 14-SEP-1998; 98WO-US019177.
```



```
QY 297 SerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThr 316
Db 1295 CCAGTCGATGGGCGTGGAGCGAGTGGAGCAAGTGTGAGCTGCAGCTAGTGTGCC 1354
QY 317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGln 336
Db 1355 CACTGGCGTAGCGCGAGTGCATGGCGGCCCCACCCAGAACGAGCGCGTGACTCAGC 1414
QY 337 GlyTrpAspLeuAspThrArgAsnCysThrSerAspLeuCysValHisSer----- 353
Db 1415 GGGACCGCTGCTCAGCTCTAAGAACTGCACAGATGGCGCTGTCATGCAAAATAAGAAAACT 1474
QY 354 -----AlaSerGlyProGluAspValAlaLeuTyr 363
Db 1475 CTAAGCGACCCCAACACCCACCTGCTGGAGCGCTCAGGG-----CATGGCGCGCTGTAT 1528
QY 364 ValGlyLeu---IleAlaValAlaValCysLeuValLeuLeuLeuValLeuLeu 382
Db 1529 GCGGGGCTCGTGTGGCCATCTTCGTGGTGGTGGCAATCTCATGGCGGTGGGGTGGTG 1588
QY 383 ValTyrCysArgGlyLysGluGlyLeuAspSerAspValAlaAspSerSer---IleLeu 401
Db 1589 GTGTACCGCGCAACTCGCGTGACTTCGACACAGACATCACTGACTCATCTGCTGCCCTG 1648
QY 402 ThrSerGlyPheGlnProValSerLysProSerLysAlaAspAsnProHisLeuLeu 421
Db 1649 ACTGTGGTTTCCACCCGCTCACTTTAGACGGCAAGGCCAGCAACCCGCGAGTCTCTA 1708
QY 422 -----ThrIleGlnProAspLeuSerThrThr---ThrTyrGlnGlySerLeuCys 438
Db 1709 CACCCCTCTGTGCTCTCACTGACAGCGCGCGGCATCTACCGCGGACCCGGTGTAT 1768
QY 439 ProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerPro 458
Db 1769 GGCCTGCGAGACTCC---ACCGACAAATCCCGATACCAACTCTCTCTGCTGGACCCC 1825
QY 459 LeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThr----- 472
Db 1826 TTACCCAGCGCTTAAGTCAAGTCTACAGCTCCAGCACACCGGGCTCTGGGCCAGGCGCTG 1885
QY 473 SerGluAlaGluPheValSerArgLeuSerThrGlnAsnTyr----- 487
Db 1886 GCAGATGGGGCTGACCTGTGGGGGTCTTGGCGCTGGGCACATACCTAGCGATTTCGCC 1945
QY 488 -----PheArgSer----- 490
Db 1946 CGGGACACCACTTCTGCACTGCGCAGCGCGAGCGCTCGGTTCACGAGCTCTTGGGC 2005
QY 491 LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeu 510
Db 2006 CTGCCCCGAGAGCCAGGAGACGCTCAGCGGCACCTTTGGCTGCTGGGTGGGAGGCTC 2065
QY 511 MetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLys 530
Db 2066 AGCATCCCCGGGCACAGGGGTGAGTGTGTGGTGGCCCAATGGAGCCATTCCCGAGGCGAAG 2125
QY 531 IleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProIleAla--- 549
Db 2126 TTCTAGAGATGATCTACTCATCAACAAGCAGAAAGTACC---CTCCCGCTTTTCAGAA 2182
QY 550 GlyCysGlnThrLeuLeuSerProIleValSerCysGlyProGlyValLeuLeuThr 569
Db 2183 GGGACCCAGACAGTATTGAGCCCTCGGTGAGCTGTGGAGCCACAGGCGCTCTGCTGTGCC 2242
QY 570 ArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeu 589
Db 2243 CGCCCGCTCATCTCACCATGCCCTCAGTGTGGAGTCAAGTCAAGTCAAGTCAAGTCAAGT 2302
QY 590 ArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGlu 609
Db 2303 CAGCTCAAGACCGGCGCCAGGCGCCACTGGAG---GAGGTGTGACCTGGATGAG 2359
QY 610 GluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThr 629
```

```
Db 2360 GAGACCTCAACACACACCCTGCTACTCCAGCTGGAGCCCGAGGCCCTGTACATCTCTGCTG 2419
QY 630 GluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaLysArg 649
Db 2420 GACCAGCTGGGCACCTACCTGTGTTCACGGCGAGTCTATTCCCGCTCAGCAGTCAAGCGG 2479
QY 650 LeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgVal 669
Db 2480 CTCACGCTGGCGCTCTTCGCCCGCCCTCTGCACCTCCCTGGAGTACAGCTCCCGGTC 2539
QY 670 TyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeu 689
Db 2540 TACTGCTGGAGCACAGCGCTTAGCAGCTGAAGAGGTCTGGAGCTGGAGCGACTCTG 2599
QY 690 GlyGlyGlnLeuLeuLeuGlnProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709
Db 2600 GCGGATACCTTGGTGGAGGAGCGAAACCGCTAACTGTTCAAGACAGTATTACCAACCTG 2659
QY 710 ArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyr 729
Db 2660 CGCCTCTCCCTCCATGACCTCCCATGCCATTGGAGGAGCAAGCTGCTGGCCAAATAC 2719
QY 730 GlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPhe 749
Db 2720 CAGGAGATCCCTTCTATCAGATTTGGAGTGGCAGCCAGAGGCCCTCCCTGACATTC 2779
QY 750 ThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGln 769
Db 2780 ACCCTGGAGAGCACAGCTTGGCTCCACAGAGCTCACCTGCAAGATCTGCTGGCGCAA 2839
QY 770 ValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPhe 789
Db 2840 GTGGAAGGAGGAGGCGCAGATATTCCAGCTGCATACCACTCTGGCA---GAGACACCTGCT 2896
QY 790 AlaGluLeuLeuAlaLeuGluSerGluAlaGly-----ValProAlaLeuValGlyPro 807
Db 2897 GGCTCCCTGGACACTCTCTGCTCTGCCCTGGCAGCAGCTGTCCACCCAGCTGGAGCT 2956
QY 808 SerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProPro 827
Db 2957 TATGCTTCAAGATCCCACTGTCATCCGCGCAGAGATATGACACAGCTAGATGCCCCC 3016
QY 828 CysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeu 847
Db 3017 AACTCAGCGGGCAATGACTGGCGGATGTAGCACAGAAGCTCTCTATGACCGGTACTG 3076
QY 848 SerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArg 867
Db 3077 AATTACTTTGCGCACAAGCGAGCCCGGCTGGTGTATCTCTGGACCTCTGGGAAGCTCTG 3136
QY 868 HisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnPro 887
Db 3137 CAGCAGGACGATGGGACCTCAACAGCTGGCGAGTGTCTTGGAGGAGATGGCAAGAGT 3196
QY 888 AspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 3197 GAGATGCTGGTGGTGTGGCCACCGGAGGACTGC 3232
```

Search completed: October 6, 2004, 18:38:43

Job time : 971 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 6, 2004, 18:19:39 ; Search time 8838 Seconds
(without alignments)
4408.847 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRPGIWPALLGIVLAAML.....AVAGLGQPDAGLFTVSEAC 899

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-O=/cn2.1/USPTO.spool/US09970944/runat_05102004.112010.1626/app.query.fasta_1.1095
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09970944 @CGN 1.1 5932 @runat_05102004.112010.1626 -NCPV=6 -ICPV=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sv:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	4787	100.0	2881	6	AX527916 Sequence
2	4698.5	98.2	2752	6	AX449572 Sequence
3	4688.5	97.9	2697	6	AX451652 Sequence
4	4592.5	95.9	3992	10	MMU487852 Sequence
5	4545.5	95.0	2697	6	AX268596 Sequence
6	4545.5	95.0	2697	10	RNU87305 Sequence
7	4545.5	95.0	3014	6	BD057524 Netrin re
8	4379.5	91.5	3580	6	AX367094 Sequence
9	4278.5	89.4	3844	10	BC058084 Mus muscu
10	3024.5	63.2	4294	10	AK122575 Mus muscu
11	2819	58.9	2688	9	BC009333 Homo sapi
12	2705.5	56.5	9299	10	MMU72634 Sequence
13	2694	56.3	1787	6	BD057525 Netrin re
14	2688.5	56.2	9328	10	AB118026 Rattus no
15	2681	56.0	200968	2	AC123700 Mus muscu
16	2679.5	56.0	2962	5	AY187310 Gallus ga
17	2673.5	55.8	3646	9	AF055634 Homo sapi
18	2570	53.7	2832	5	AY099459 Xenopus l
19	2513.5	52.5	166534	2	AC034209 Homo sapi
20	2505.5	52.3	3672	10	BC057560 Mus muscu
21	2505.5	52.3	3672	10	BC048162 Mus muscu
22	2499	52.2	3788	10	MMU487853 Sequence
23	2497	52.2	2838	10	RNU87306 Sequence
24	2490.5	52.0	2995	6	AX497288 Rattus norv
25	2486.5	51.9	3770	9	AY126437 Homo sapi
26	2484	51.9	2860	6	AX686445 Sequence
27	2484	51.9	2860	6	AX686447 Sequence
28	2479	51.8	3884	6	AX464012 Sequence
29	2479	51.8	3884	9	AY358351 Homo sapi
30	2479	51.8	3933	9	AB128132 Homo sapi
31	2479	51.8	3935	9	AB096256 Homo sapi
32	2472	51.6	2895	6	AX512381 Sequence
33	2463	51.5	2831	6	BD057526 Netrin re
34	2444.5	51.1	168168	2	AC012283 Homo sapi
35	2381	49.7	121282	9	AC027318 Homo sapi
36	2276.5	47.6	9700	6	AX054976 Sequence
37	2194.5	45.8	231407	2	AC139592 Rattus no
38	2148	44.9	2868	6	AX800719 Sequence
39	2146	44.8	3252	9	AY358147 Homo sapi
40	2144.5	44.8	3925	10	MMU487854 Mus muscu
41	2140.5	44.7	2661	6	AX800717 Sequence
42	2136	44.6	7002	6	BD162193 Novel unc
43	2136	44.6	7002	9	AB055056 Homo sapi
44	1646	34.4	6495	9	AB075856 Homo sapi
45	1646	34.4	6522	6	BD186447 Novel gen

ALIGNMENTS

RESULT 1

AX527916
LOCUS AX527916 2881 bp DNA linear PAT 21-NOV-2002
DEFINITION Sequence 1 from Patent WO0229038.
ACCESSION AX527916
VERSION AX527916.1 GI:25172359
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Herrmann, J.L., Rastelli, L. and Shinkets, R.A.
TITLE Novel proteins and nucleic acids encoding same and antibodies
directed against these proteins
JOURNAL Patent: WO 0229038-A 1 11-APR-2002;
Curagen Corporation (US)
FEATURES
source
1..2881
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6.13e-295 Length: 2881
Score: 4787.00 Matches: 899
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-970-944-2 (1-899) x AX527916 (1-2881)
QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyLeValLeuAlaTrpLeu 20
Db 87 ATGCCCGTCCGGCCGCGCTGTGGCCAGCGCTCTGGGCATAGTCTCGCCGCTTGCTC 146
QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db 147 CGCGGCTCGGGTGGCCAGCAGAGTGCACCGTGGCCAAACAGTGGCTGTGGCAACCCG 206
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTrileValLysAsnLysPro 60
Db 207 GACCTGCTTCCCACTTCTCTGGTGGAGCCGAGAGTGTACATCTGTCAGAACAGCCA 266
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db 267 GTGCTGCTTGTGTCAAGCGCGTCCCGCCACGCGAGATCTTCTCAAGTGCACCGGGAG 326
QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db 327 TGGGTGGCCAGGTGGACCGATGATCGAGCGACGACAGCGGAGCAGTGGTGGAGCCG 386
QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db 387 ACCATGAGGTCCCGATTATGCTCTCAAGCAGCAGAGTTCGNGAAGTGTTCGGGCTGGAG 446
QY 121 GluTrpTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db 447 GAATACTGGTCCAGTGGCTGGCATGGAGTCTCTCGGCGACACCAAGAGTGCAGAGGCC 506
QY 141 TyrlleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db 507 TACATCCGATAGCCAGATTCCGGAAGACTTCAGCAGAGAGCCGCTGGCCAGAGAGGTG 566
QY 161 SerLeuGluGlnGlyLeValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db 567 TCCCTGGAGCAGGGCATCGTGTGCTGCCCTGCCCTCCACCGAGGAGCATCCCTCCAGCCGAG 626
QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTrile 200
Db 627 GTGAGTGGCTCCGGAACGAGGACCTGGTGGACCGCTCCCTGGAGCCCAATGTATACATC 686

201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTrpTyr 220
Db 687 ACCGGGAGACACAGCTGTGTGGACAGGCGCGCTTGTGTGACAGGCCCACTACACC 746
QY 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaValIleValTyr 240
Db 747 TGGCTGGCCAAAGAACATCGTGGCAGCTCGCGCAGCGCTCCCGCTGTGTATCGTCTAC 806
QY 241 ValAsnGlyGlyTrpSerThrTrpGluTrpSerValCysSerAlaSerCysGlyArg 260
Db 807 GTGAACCGTGGGTGGTGGACCGAGTGTGTGGTGTGGACGCGCAGCTGTGGGCGC 866
QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
Db 867 GGTGTGACAGAACGAGGAGCGGAGCTGCACCAACCGCGGCTCTCAACGGGGGCGTTTC 926
QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
Db 927 TGTAGGGGCGAGAAATGTCCATGACCGACCGTCTCTCTCTGTGTGTGTGGACGGC 986
QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db 987 AGCTGAGCCCGTGGAGCAAGTGTGGCGCTGTGGGCTGGACTGACCCACTGGCGGAGC 1046
QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db 1047 CGTGTGTGTCTGTACCCAGCACCCCGACGAGGGGAGGAGTGGCAGGGCACTGACCTG 1106
QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db 1107 GACACCGCGAACTGTATACAGTGTGTGTACACAGTGTCTTGTGGCCCTGGAGGAGTG 1166
QY 361 AlaLeuTrpValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
Db 1167 GCCTCTATGTGGCTCTATCGCGCGTGGCGCTGTGGCTGTCTGTCTGTCTGTCTGTCT 1226
QY 381 IleLeuValTrpCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerile 400
Db 1227 ATCTCTGTATTATTCGCGAAGAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db 1287 CTCACCTCAGCTTCCAGCCCGTCCAGCATCAAGCCAGCAAGCAGCAACCCCATCTG 1346
QY 421 LeuThrIleGlnProAspLeuSerThrThrThrThrGlnGlySerLeuCysProArg 440
Db 1347 CTCACCATCCAGCGGACCTCAGCACCCACCCCTACCCAGGCGAGTCTGTGTCCCGG 1406
QY 441 GlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGly 460
Db 1407 CAGGATGGGCGCCAGCCCAAGTTCAGCTCACCAGTGGCACCTGTCTCAGCCCTGGGT 1466
QY 461 GlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSer 480
Db 1467 GGGCGCGCCACACACTGTGCCACCACTCTCCACCTCTCAGGCGCGAGGATTCGTCTCC 1526
QY 481 ArgLeuSerThrGlnAsnTrpPheArgSerLeuProArgGlyThrSerAsnMetThrTyr 500
Db 1527 CGCTCTCCACCCAGAACTACTTCCGCTCTCCCGCGAGCCACCAACATGACCTAT 1586
QY 501 GlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeu 520
Db 1587 GGGACCTTCAACTCTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGGCTCTCTC 1646
QY 521 IleProProAspAlaIleProArgGlyLysIleTrpGluIleTrpLeuThrLeuHisLys 540
Db 1647 ATCCCCCGCATGCGCATACCCCGAGGAGATCTATGAGATCTACCTCACCTGCACAG 1706
QY 541 ProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSer 560
Db 1707 CCGGAAGACGTGAGTTGGCCCTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1766
QY 561 CysGlyProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGly 580

```
Db 1767 TGTGACCCCTGGCGTCTGCTACCCGGCAGTCATCCTGGCTATGACCACTGTGGG 1826
Qy 581 GluProSerProAspSerTrpSerLeuArdLeuLysGlnSerCysGluGlySerTrp 600
Db 1827 GAGCCAGCCCTGACAGCTGGAGCCCTGGCCCTCAAAAAGCAGTCGTGCGAGGCGAGCTGG 1886
Qy 601 GluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeu 620
Db 1887 GAGCAGGATGTGCTGCACCTGGCGAGGAGGGCCCTCCCACTCTACTACTGCGAGCTG 1946
Qy 621 GluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGlu 640
Db 1947 GAGGCCAGTGGCTGCTAGCTTTCACCGAGCAGCTGGGCCCTTTGCCCTGGTGGAGAG 2006
Qy 641 AlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCys 660
Db 2007 GCCCTCAGCTGGCTGGCGCCAGCGCTCAAGCTGCTTCTGTTGGCGCGGTGGCTGCG 2066
Qy 661 ThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLys 680
Db 2067 ACCTCCCTCGAGTACAACATCCGGTCTACTGCTGCTGATGACACCCAGGATGCACTCAAG 2126
Qy 681 GluValValGlnLeuGluLysGlnLeuGlyGlnLeuLeuGlnGluProArgValLeu 700
Db 2127 GAGTGTGTGAGCTGGAGAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCTG 2186
Qy 701 HisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeu 720
Db 2187 CACTTCAAGACAGTACCAACCTGGCGCTATCCATCCACGATGCGCCAGCTCCCTG 2246
Qy 721 TrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGly 740
Db 2247 TCGAAGAGTAAAGCTCCTGTGTCAGTACCAGGAGATCCCTTTATCATCTGGAATGGC 2306
Qy 741 ThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAsp 760
Db 2307 ACGCAGCGGTACTTGCACCTGACCTTCCAGCTGGAGCGTGTGACGCCAGCACTAGTGAC 2366
Qy 761 LeuAlaCysLysLeuTyrValTyrGlnValGluGlyAspGlyGlnSerPheSerIleAsn 780
Db 2367 CTGGCTGCAAGCTGTGGGTGTGGAGTGGAGGGGAGCGGAGAGCTTTCAGCATCAAC 2426
Qy 781 PheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly 800
Db 2427 TTCAACATCAACCAAGGACACAAGGTTTGTGAGCTGCTGCTGTGGAGATGAAGCGGG 2486
Qy 801 ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIle 820
Db 2487 GTCCCAAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAAGATA 2546
Qy 821 IleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLys 840
Db 2547 ATTTCCAGCTGGACCAACCTGTAGCGGGGTGGAGCTGGCGGACTCTGGGCCCAAGAA 2606
Qy 841 LeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIle 860
Db 2607 CTCACCTGGACAGCCATCTCAGCTTCTTGGCTTCCAAAGCCAGCCAGCCAGCCATGATC 2666
Qy 861 LeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 880
Db 2667 CTCACCTGTGGAGGCGGCGCATCTTCCCAACGGCAACCTCAGCCAGCTGGCTGGAGCA 2726
Qy 881 ValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 2727 GTGGCTGGAGTGGCGCAGCAGCAGTGGCCCTTTCACAGTGTGCGAGGCTGAGTGC 2783

RESULT 2
AX449572 LOCUS
DEFINITION Sequence 1 from Patent WO0210216.
ACCESSION AX449572
VERSION AX449572.1 GI:21698195
```

```
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Padigaru,M., Mezes,P., Mishra,V., Burgess,C., Casman,S.,
Grosse,W.M., Alsbrook,J.P., Iepley,D.M., Gerlach,V.L.,
Macdougall,J.R. and Smithson,G.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 0210216-A 1 07-FEB-2002;
FEATURES Curagen Corporation (US)
source Location/Qualifiers
1..2752
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 2,45e-289 Length: 2752
Pred. No.: 4698.50 Matches: 888
Score: 98.89% Conservative: 2
Percent Similarity: 98.67% Mismatches: 7
Best Local Similarity: 98.15% Indels: 3
Query Match: 6 Gaps: 3
DB:

US-09-970-944-2 (1-899) x AX449572 (1-2752)
Qy 1 MetAlaValArgProGlyLeuTyrProAlaLeuLeuGlyIleValLeuAlaAaTrpLeu 20
Db 46 ATGGCCGTCGCGCCCGCCCTGTGCCAGCGCTCTTGGGCATAGTCTCGCGCTGTGCTC 105
Qy 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db 106 CGCGGCTCGGCTGCCAGCAGAGTGCACCGTGGCCACCCAGGCTGCTGGTGCACCCG 165
Qy 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db 166 GACTCTGTCCTCCACATTCCTGCTGGAGCCGAGGATGTATCATCTCTTCAAGTCAACG 225
Qy 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db 226 GTGCTGCTGTGTGCAAGGCCGTGCCCGCCAGCATCTTCTTCAAGTCAACGCGGAG 285
Qy 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db 286 TGGGTGCGCCAGCTGGACCGATGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCC 345
Qy 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db 346 ACCATCGAGTCCGCAATTATGTCTCAAGGCAGCAGTCTCGAGAGAGGCTGTTGGGCTGG 405
Qy 121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThryLysSerGlnLysAla 140
Db 406 GAATATCTGTGTCAGTGTGCTGGCATGGAGCTCTCGGCGCAGCAGTCTGAGAAAGGTC 465
Qy 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGlnProLeuAlaLysGluVal 160
Db 466 TACATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 525
Qy 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProAlaGlu 180
Db 526 TCCCTGGAGCAGGCGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 585
Qy 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db 586 GTGGAGTGGCTCCGAACAGGAGCCTGTGTGGAGCCCGCTCCCTGAGACCCCAATGATATC 645
Qy 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
Db 646 ACGCGGAGCACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 705
```

221 QY CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240
DB TGCCTGGCCAGAACAATCGTGGCAGCGTCCGCGAGCGCTCCGCTGTCATCGTCTAC 765
241 QY ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
DB GTGAACCGTGGGTGGTCCAGCTGACCGAGTGGTCCGCTGTCAGCGCCAGCTGTGGCGC 825
261 QY GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
DB GGCTGGCAGAAACGGAGCGAGCTGCACCAACCCGCGCTCTCAACGGGGCGCTTTC 885
281 QY CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
DB TGTGAGGGGAGAGTGC---CAGAAACACCGCTGCGCCACCCCTGTCGCCAGTAGACGGC 942
301 QY SerTrpSerProTrpSerLysLeuTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
DB AGCTGGAGCCGTTGGAGCAAGTGTGCGCTGTGGCTGTGGACTGCACCCACTGCGCGAGC 1002
321 QY ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
DB CGTGAGTGTCTGACCCAGCACCCCGCAACGGAGGGAGAGTGCCAGGGCACTGACCTG 1062
341 QY AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
DB GACACCGCAACTGTACGAGTACCTCTGTGTACACAGTGTCTTGGCCCTGGAGCGTG 1122
361 QY AlaLeuTrpValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
DB GCCCTCTATGTGGCCCTATCGCCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1182
381 QY IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
DB ATCTCTGTTATTCCCGAAGAGAGGGGTGGACTCAGACTGTGGTGTGACTGTGCTGCTGCT 1242
401 QY LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
DB CTCACCTCAGGCTTCCAGCCCGTCAAGCATCAAGCCCGCAGCAAGCAGCAACCCCACTG 1302
421 QY LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThrThrThr 439
DB CTCACCTCAGGCTTCCAGCCCGTCAAGCATCAAGCCCGCAGCAAGCAGCAACCCCACTG 1362
440 QY ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
DB CGGCAGGATGGCCAGCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCTG 1422
460 QY GlyGlyGlyArgHisThrLeuHisSerSerProThrSerGluAlaGluGluPheVal 479
DB GGTGGCGCCGCCACACACTGCACACAGCTCTCCCACTCTGAGGCGGAGAGTTCGTC 1482
480 QY SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
DB TCCGCTCTCCACCCAGAACTACTTCGCTCCCTGCCCGCCGAGCCAGCAACATGACC 1542
500 QY TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
DB TATGGACCTTCAACTTCTCGGGGGCGGTGATGATCCCTTAATACAGGTATCAGCCTC 1602
520 QY LeuIleProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
DB CTCTATCCCCCAGATGCCATACCCCGAGGAGAGATCTATGAGATCTACTCTACGCTGCAC 1662
540 QY LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
DB AAGCCGAGAGAGTGAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1722
560 QY SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
DB AGCTGTGGACCCCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1782
580 QY GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599

1783 DB GGGAGCCCAAGCCCTGACAGCTGGAGCTCGCCTCAAAAGCAGTCTGTCGGAGGGCAGC 1842
600 QY TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
DB TGGGAG---GATGTGTGACCTGGGCGAGAGCGCCCTCCACCTCTACTACTGCGCAG 1859
620 QY LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
DB CTGGAGGCCAGTGCCTGTACGTCTTACCAGAGCAGCTGGGCGCTTGGCTTGGTGGGA 1959
640 QY GluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla 659
DB GAGGCCCTCAGCGTGGTGGCGCAAGCGCTCAAGCTGCTTCTGTTGGCGCGTGGCC 2019
660 QY CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
DB TGCACCTCCCTCGAGTACAACTCCGGGTCTACTGCTGCTGATGACACCCAGATGCACTC 2079
680 QY LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnProArgVal 699
DB AAGAGGTGTGTGAGCTGGAGAGCAGCTGGGGGAGCAGCTGATCCAGGAGCCACGGGTC 2139
700 QY LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
DB CTGCACCTTCAAGGACAGTTACCACAACCTGCGCTATCCATCCACGATGTGCCAGCTCC 2199
720 QY LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
DB CTGTGGAAGAGTAACTCTTGTGAGTACAGAGATCCCTTTTATCATCATCTGGAAT 2259
740 QY GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
DB GGCACGAGCGGTACTTGTGACTGCACCTTACCCTGGAGCGTGTGACCCCGCAGCACTAGT 2319
760 QY AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
DB GACCTGGCTGCAAGCTGTGGTGTGGCAGGTGGAGGGCGACGGCGAGCTTCAGCATC 2379
780 QY AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
DB AACTTCAACATCAACAAGGACAAAGTGTGCTGAGCTGTGGCTGTGGAGGTGAAGCG 2439
800 QY GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
DB GGGGTCCAGCCCTGTGGGCCCCAGTGCCTTCAAGATCCCTTCCCTCATTCGGCAGAG 2499
820 QY IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
DB ATATTTCCAGCCTGGACCCACCTGTAGGGGGGTGCGGACTGCGGACTCTGCGCCAG 2559
840 QY LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
DB AAATCCACCTGGACGCCATCTCAGCTTCTTGTCCCTCAAGCCCGAGCCCGCAGCCATG 2619
860 QY IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
DB ATCTTCAACCTGTGGAGCGCGGACCTTCCCACGCAACCTCAGCCAGCTGGCTGCA 2679
880 QY AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
DB CGATGGTGTGAGTGGCCAGCCAGACGCTGGCTTCTTACAGTGTGCGAGGCTGAGTGC 2739

RESULT 3

AX451652

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AX451652
Sequence 1 from Patent WO0233080.
AX451652.1 GI:21698587

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

linear

DNA

2697 bp

PAT 03-JUL-2002

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 Koehler, R.H.
 Regulation of human netrin binding membrane receptor unc5h-1
 Patent: WO 0233080-A 1 25-APR-2002;
 Bayer Aktiengesellschaft (DE)
 Location/Qualifiers
 1..2697
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

FEATURES
 source

ORIGIN
 Alignment Scores:
 Pred. No.: 1.03e-288 Length: 2697
 Score: 4688.50 Matches: 986
 Percent Similarity: 98.78% Conservative: 3
 Best Local Similarity: 98.44% Mismatches: 8
 Query Match: 97.94% Indels: 3
 DB: 6 Gaps: 3

US-09-970-944-2 (1-899) x AX451652 (1-2697)

QY	1	MetAlaValArgProGlyLeuTyrProAlaLeuLeuGlyIleValLeuAlaAlaTyrLeu	20
Db	1	ATGGCCCTCGGGCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCTCGCGCTTGGCTC	60
QY	21	ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro	40
Db	61	CGCGGCTCGGGTGCCAGCAGTGCCACCGGTGGCCACCCAGTGCCTGGTGGCAACCCG	120
QY	41	AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro	60
Db	121	GACCTGCTTCCCACTTCTGGTGAGCCCGAGATGTGTACATCGTCAAGAACAGCCA	180
QY	61	ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu	80
Db	181	GTGCTGCTGTGTGCAAGGCGTGCCTGGCCAGCAGATCTTCTCAAGTGCACGGGGAG	240
QY	81	TyrValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro	100
Db	241	TGGGTGGCCAGGTGGACACCGATGATCGAGCGCAGCAGACGCGGAGCAGTGGGCTGCC	300
QY	101	ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu	120
Db	301	ACCATGAGTCCGCAATTAATGTTCTCAAGCAGCAGGTCCGAGAGGTGTTCGGGCTGGAG	360
QY	121	GluTyrTrpCysGlnCysValAlaTyrSerSerSerGlyThrThrLysSerGlnLysAla	140
Db	361	GAATACTGGTGCAGTSCGTGGCATGGAGTCTCTCGGGCCACCACCAAGATCAGAGGCC	420
QY	141	TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal	160
Db	421	TACATCGCATACGCTATTTCGCAAGAACTTCGAGCAGAGCGCGCTGGCCAAAGAGGTG	480
QY	161	SerLeuGluGlnGlyIleValLeuProCysArgProGluGlyIleProProAlaGlu	180
Db	481	TCCTTGAGCAGGGCATCGTGCTGCCCTCGCTCCACCGAGGGGCATCCCTCCAGCCGAG	540
QY	181	ValGluTyrLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle	200
Db	541	GTGGAGTGGCTCGGAACAGAGACCTTGGTGGACCGCTCCGTCCTGGACCCCAATGATCATC	600
QY	201	ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr	220
Db	601	ACGCGGAGCAGACCTGTGGTGGCAGACGGCCGCTTGTGACACGGCCAACTACACC	660
QY	221	CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr	240
Db	661	TGGGTGGCCAAAGACATCGTGGACGTCGCGCAGCGGCTCCGCTGCTGTCTACTCGCTAC	720
QY	241	ValAsnGlyClyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg	260

Db	721	GTGAACGCTGGTGTGCACCTGGACCGAGTGGTCCGCTGTGCAGGCCAGCTGTGGGCGC	780
QY	261	GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe	280
Db	781	GGCTGGCAGAAACAGAGCGGAGCTGCACCAACCGCGCCTCTCAACGGGGCGCTTTC	840
QY	281	CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly	300
Db	841	TGTGAGGGGCGAGAATGTCTCAGAAAAACACCTCGCGCCACCTGTGCCAGTGGACGCGC	897
QY	301	SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer	320
Db	898	AGCTGGAGCCCGTGGAGCAAGTGTGGCCCTGTGGCTGGACTGCACCATCTGGCGGAGC	957
QY	321	ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGlnCysGlnGlyThrAspLeu	340
Db	958	CGTGAGTGTCTGACCCAGCACCAGCAACGAGGAGGAGTGCAGGCGACTGACCTG	1017
QY	341	AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal	360
Db	1018	GACACCCGCAACTGTACAGTACCTCTGTGTACACACTGTCTTGGCCCTGAGGACGCTG	1077
QY	361	AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu	380
Db	1078	GCCCTCTATGTGGCCCTCATCGCGTGGCCGTCTGCCCTGGTCTGCTGCTGTGCTC	1137
QY	381	IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle	400
Db	1138	ATCTCTGTTATTTCGCGAAGAGAGGGGTGACTCAGATGATGTGGCTGACTCTGCTCAT	1197
QY	401	LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu	420
Db	1198	CTACCTCAGCTTCCAGCCCGTTCAGCATCAAGCCAGCAAGCAGACCAACCCCATCTG	1257
QY	421	LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrGlnGlySerLeuCysPro	439
Db	1258	CTCACCATCCAGCCGAGCTCAGCACCCACCAACCACTTACAGGCGAGTCTCTGTGCC	1317
QY	440	ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu	459
Db	1318	CGGAGGATGGGCCAGCCCAAGTTCAGCTCACCATAATGGGCACCTGCTCAGCCCCCTG	1377
QY	460	GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal	479
Db	1378	GGTGGGCGCGCCACACACTGCACACAGCTCTCCACCTCTGAGGCGGAGGTTGCTC	1437
QY	480	SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr	499
Db	1438	TCCCGCTCTCCACCCAGCAACTACTTCCGCTCCCTGCCCGAGGCGCACCAACATGACC	1497
QY	500	TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu	519
Db	1498	TATGGACCTTCAACTTCTCGGGGCGCGCTGATGATCCCTAATACAGGAATACGCTC	1557
QY	520	LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrIleThrLeuHis	539
Db	1558	CTCATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTACGCTGCAC	1617
QY	540	LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal	559
Db	1618	AGGCGGAGAGAGCTGAGGTGGCTTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	1677
QY	560	SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys	579
Db	1678	AGCTGTGGACCCCTGGCGTCTGCTCACCAGGCGAGTCTCTGCTGTATGGACCACTGT	1737
QY	580	GlyGluProSerProAspSerTrpSerLeuArgLeuLysGlnSerCysGluGlySer	599
Db	1738	GGGAGAGCCAGCCCTGACAGCTGGAGCCTCGCCTCAAAAAGCAGCTGTGCGAGGCGAG	1797
QY	600	TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln	619
Db	1798	TGGGAG---GATGTCTGCACCTTGGCGAGGAGGCGCCCTCCGCTCTACTACTGCGCAG	1854

```
QY 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db 1855 CTGGAGCCAGTGGCTCTACGCTTCCACGAGCAGCTGGCGCTTTGGCCCTGGTGGGA 1914
QY 640 GluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla 659
Db 1915 GAGGCCCTCAGCGTGGTGGCGGCAAGCGCTCAAGCTCTTCTTTTGGCGCGTGGCC 1974
QY 660 CysThrSerLeuGluTyrAsnLysArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db 1975 TGACCTCCCTCGAGTACACATCCGGGTCTACTGCTCGATGACACCCAGATGCACTC 2034
QY 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuLysGlnLeuProArgVal 699
Db 2035 AAGGAGGTGGTGGAGTGGAGAGCAGCTGGGGGACAGCTGATCCAGGAGCCACCGGTC 2094
QY 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db 2095 CTGCACCTTCAAGGACAGTTACCAACCTGGCGCTATCCATCCAGATGGCCAGCTCC 2154
QY 720 LeuTyrLysSerLysLeuLeuValSerTyrGlnGluLysProPheTyrHisIleTyrAsn 739
Db 2155 CTGTGGAGAGTAAGCTCTTGTGAGCTACAGGAGATCCCTTTTATCATCATCTGGAAT 2214
QY 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db 2215 GGCACGAGCGGTACTTGCATCGACCTTACCTTGGAGCGTGTGAGCCCGAGCACTAGT 2274
QY 760 AspLeuAlaCysLysLeuTyrValTyrGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db 2275 GACCTGGCGCTCAAGCTGTGGGTGTGGAGTGGAGGCGGAGGAGCTTTCAGCATC 2334
QY 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db 2335 AACTTCAACATCACCAAGGACACAAAGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
QY 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuLysArgGlnLys 819
Db 2395 GGGGTCCAGCGCTGGTGGCGCCAGTGGCTTCAAGATCCCTTCTCATTCGGCAGAG 2454
QY 820 IleIleSerSerLeuAspProCysArgArgGlyAlaAspTyrArgThrLeuAlaGln 839
Db 2455 ATAATTTCCAGCTTGAGCCACCTCTGAGGCGGGTGGCGACTGCGGACTCTGGCCAG 2514
QY 840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db 2515 AAATCCACCTGGACAGCACTCTCAGCTTCTTGGCTCCAGGCCAGCCACAGCCATG 2574
QY 860 IleLeuAsnLeuTyrGluAlaArgHisPheProAsnGlyAsnLeuSerGluLeuAlaLa 879
Db 2575 ATCTCTCAACCTGTGGAGCGCGGCACTTCCCCAACGGCAACCTCAGCAGCTGGCTGCA 2634
QY 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 2635 GCAGTGGCTGGACTGGGCGAGCCAGCAGCTGGCTTCTTCACTGCTCGAGGCTGAGTGC 2694

RESULT 4
MMU487852 3992 bp mRNA linear ROD 24-SEP-2002
LOCUS Mus musculus mRNA for netrin receptor Unc5h1 (Unc5h1 gene).
DEFINITION
ACCESSION AJ487852
VERSION AJ487852.1 GI:22035783
KEYWORDS netrin receptor Unc5h1; Unc5h1 gene.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1
REFERENCE
AUTHORS Engelkamp,D.
TITLE Cloning of three mouse Unc5 genes and their expression patterns at
mid-gestation
JOURNAL Mech. Dev. 118 (1-2), 191-197 (2002)
```

```
22239710 MEDLINE
12351186 PUBMED
2 (bases 1 to 3992) REFERENCE
Engelkamp,D. AUTHORS
Direct Submission TITLE
Submitted (15-MAY-2002) Neuroanatomy, Max Planck Institute for
Brain Research, Deutschoordenstrasse 46, Frankfurt 60528, GERMANY
JOURNAL
FEATURES
Location/Qualifiers
source
1..3992
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..3992
/gene="Unc5h1"
232..2928
/gene="Unc5h1"
/codon_start=1
/product="netrin receptor Unc5h1"
/protein_id="CAD32250.1"
/db_xref="GI:22035784"
/db_xref="GOA:Q8K1S4"
/db_xref="SPTRMBL:Q8K1S4"
/translation="MAVRPGLWPALLGIVLTAMLRGSGAQSATVANPVGANPDLLP
HFIVEPDVYIVNKPVLVLYKAVPATQIEFKNGEWMVRQVDHVIERTDSSGLPTM
EVRINVRQOVKEVFGLEYQCQVAMSSSGTTKSQKAYIRIAYLRNFQEQELAKEV
SLQGIIVLPCRPEGIPPAEVEIRNEDLVDPSLDPNVITREHSLVVRQARLADTAN
YTCVAKNIVARRBSASAALVYNGGWSWTWESVCSACSGRWKRSCTNPAPLN
GGAFCEGQVQKTACATLCPVDGSWSPKWSACGLDCTHWRSECDPAPRNGERC
DADLDPTNCTSDLCHTSSGDEVDALYIQLVAVAVCLILILLVLVLVLIYCRKKEGLS
DVADSSLSFGFVSIPKSPADNPHELLTIOPLSTTTTQSSLCPRQDGPSPKFLQ
SNGHLSPGSGRHTLHSSPTSEAEFVSRSLSTQNYFRSLPRGTSNMYGTNFPFGG
RLMPTNGTSLILPPDAIPRGKIVEIYTLHKPEDVRLPLAGCOTLLSPVCSGPPGV
LLRFPVILAMDHGCEPSDSURLKKQSCGSEWEDVLHGESPSHLVYCDLHAKEVV
YVTEQGRFALVGEALSVATKRLRLLLPAPVACTSLEYNIRVYCLDTHDLKEVV
QLKQJGGQIIEPRVLHFKDSVHNRLSHDVPSSLKWSKLIVSYQEIPIYHWNGT
QVVLGFTFLIRVNASTDACKWQVWQVGDGOSFNINFTKDTFPAEMLALESG
GGPALVGFSAFKIPFLIRQKIITSLDPPCFSGADWRTLAQKLHLSLFFASKESPT
AMTILNWEARHFFNPNGLQLAAAVAGLQPDAGLFTVSEAC"
```

Alignment Scores:

Pred. No.:	2.19e-282	Length:	3992
Score:	4592.50	Matches:	858
Percent Similarity:	97.67%	Conservative:	21
Best Local Similarity:	95.33%	Mismatches:	18
Query Match:	95.94%	Indels:	3
DB:	10	Gaps:	3

US-09-970-944-2 (1-899) x MMU487852 (1-3992)

QY 1 MetAlaValArgProGlyLeuTyrProAlaLeuLeuGlyIleValLeuAlaAlaTyrLeu 20

Db 232 ATGGCGGTCGCGCGCGCTGTGGCCAGCGCTCTGGGCATAGTCTCTACCTGCGCTGTT 291

QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40

Db 292 CGTGGTTCGGTCCGCCAGCAGAGTGCCACAGTGGCCAAACCAGTGGTGGCAACCCA 351

QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60

Db 352 GACCTGTCGCCCACTTCTCTGTAGAGCGGAGGACGTGTATCATTTGTCAAGAACAAAGCCC 411

QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80

Db 412 GTGCTGCTGGTGTGCAAGCGTGTGCCGCCACCCAGATCTTCTTCAAGTGCAACGGGAA 471

QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100

Db 472 TGGGTTCGCCAGTGCATCGATCATCGAACGACGACCTGACGGCAGCAGCGGATTGCCA 531

QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120

Db 532 ACCATGGAGTCCCGGATCAACGTATCAAGGCGAGCGAGTGGTGGTGGTGGTGGTGG 591

QY 121 GluTyrTrpCysGlnCysValAlaAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
DB 592 GAGTACTGGTGCAGCTGNGTGGCATGGAGCTCTCAGGACACACAAAAGCCAGAGGCC 651
QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
DB 652 TACATCCGGATTGCGCTATTGCGCAAGAACTTTGAGCAGGAGCGCTGGCCAAAGAAAGTG 711
QY 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
DB 712 TCATGGAGCAAGGCAATTGTGCTACCTTGTGCGCCCGCGAAGGAATCCCGCCAGCTGAG 771
QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
DB 772 GTGGAGTGGCTCCGAATGAGGACCTCGTGACCCCTCCCTCGACCCCAATGTGTACATC 831
QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
DB 832 ACACGGGAGCACAGCTAGTGTGGCGGAGCGCGCTGGCGGACACTGCCAACTACACC 891
QY 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240
DB 892 TGGTGGCCCAAGACATCGTGGCCCGTGGCGAAGCGCTCTGGCGCGCTCATTTGTTTAT 951
QY 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
DB 952 GTGAACGGTGGGTGGTGCAGCTGACCGAGTGGTCCGTCTGCAGTGCAGCTGTGGCGGT 1011
QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
DB 1012 GGTGGGAGAAACGAGCGGAGCTGCACCAACCGCGCACCTCTCAACGGGGCGCTTC 1071
QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
DB 1072 TGTGAGGGGAGAAATGTC---CAGAAACACAGCTGGCGCCACTCTGTGCCAGTGATGGG 1128
QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
DB 1129 AGCTGGAGCCCATGGAGTAAGTGGTGCAGCTCGGGGCTTGCAGTGCACCCACTGGCGGAGC 1188
QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
DB 1189 CGGGAGTGTCTCCAGCCAGCGCCCGCGCAACGGAGGTGAGGAGTGGCGGGTGTGACCTG 1248
QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
DB 1249 GACACCCGCAACTGTACCAAGTACCTCTGCTGTGCACACACCTCTTCGGGCGCCGAGGACGTG 1308
QY 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
DB 1309 GCTCTCTACATCGCCCTCGTGGCGGTGGCGGTGTGCTCATCTGTGCTGTGGTCTCTC 1368
QY 381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerIle 400
DB 1369 GTCTCATCTACTCCGCGCAAGAAAGGAGACTGGACTCAGACGTGGCTGACTCATCCATC 1428
QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
DB 1429 CTTACTCTAGGCTTCAGGCTGTGACATCAAGCCAGCAAGCAACCAATCCCACTCTG 1488
QY 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrGlnGlySerLeuCysPro 439
DB 1489 CTCACCATCAACCGGACCTCAGCACCAACACACCACTTACAGGGAGCGCTGTGTCCC 1548
QY 440 ArgGluAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
DB 1549 CGGAGGATGGACCCAGCCCAAGTTCCAGCTCTCTAATGGTCACTGTCTGACCCACTG 1608
QY 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
DB 1609 GGCAGTGGCGGCATACGCTGCACCAACAGCTCCCCACCTCTGAGGCTGAGGACTTCGTC 1668

QY 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
DB 1669 TCCGCGCTCTCCACCCAAAACACTATTCTGCCCCCGGCTACCAAGCAACATGGCC 1728
QY 500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
DB 1729 TATGGGACCTTCACTTCTCGGGGGCGGCTGATGATCCCTTAACACAGGAATACGCTC 1788
QY 520 LeuIleProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
DB 1789 CTCATACCCCGGACGCGCATCCCGGAGAAAGATCTACGAGATCTACTCTACTCTGCAC 1848
QY 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
DB 1849 AAGCCAGAAAGACGTGAGGTTGCCCTAGCTGGCTGTGAGACCTGCTGAGTCTATCGTT 1908
QY 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
DB 1909 AGCTGTGGGCCCCCAGGAGTCTGCTCACCGCGCCAGTCACTCTTGCATCGACACTGC 1968
QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
DB 1969 GGGGAGCCAGTCCCGACAGCTGGAGCTCGCCCTCAAAAAGCAGTCTCTGTGAGGGCAGC 2028
QY 600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
DB 2029 TGGAG---GACGTGTGTCACCTTGGTGGAGGAGTGGCCCTCTCATCTCTACTTGCAC 2085
QY 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
DB 2086 CTGGAGCCCGGGGCTGCTATGCTTCCACGAGCAGCTAGCCGCTTGGCCCTGGTGGGA 2145
QY 640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
DB 2146 GAGGCCCTCAGCTGGCTGCCCAACGCGCTCAGGCTCTCTGTGTGGCCCTGTGGCC 2205
QY 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
DB 2206 TGTACGTCTCAGTACAAACATCCGAGTGTACTGCTGCACGACACCCACGATGCTCTC 2265
QY 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
DB 2266 AAGGAGTGTGTGAGCTGGAGAACAGCTGGGTGGACAGCTGATCCAGGAGCCCCGTGTC 2325
QY 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
DB 2326 CTGCACTTCAAAACAGAGTTACCAACACTCCAGTCTGTCCATCCACGAGTGGCCAGCTCC 2385
QY 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
DB 2386 CTGTGAAAGAGCAAGCTCTCTGTACGTACCAAGAGATCCCTTTTACCACATCTGGAAT 2445
QY 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
DB 2446 GGCACCTCAGAGTATCTGCAGTGCACCTTACCTCGAGCGCGCTCAATCCAGCACGAGC 2505
QY 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
DB 2506 GACCTGGCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGACAGAGCTTCAACATC 2565
QY 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
DB 2566 AACTTTAAATCACTAAGGACACGAGGTTGTGAAATGCTGGCTCTGGAGAGTGAAGGG 2625
QY 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
DB 2626 GGGGTCCCAAGCCCTGGTGGGCCCCAGTGTCTTCAAGATCCCTTCTCTCATTCGGCAAAAG 2685
QY 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
DB 2686 ATCATTACCAGCTGGACCCACCTTCGAGCGGGCGCGACTGGCGAACTCTTAGCCCCAG 2745
QY 840 LysLeuHisLeuAspSerHisLeuSerPheAlaSerLysProSerProThrAlaMet 859


```

QY 500 TyrGlyThrPheAsnProGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db 1498 TACGGGAGCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACGGGATCAGCCTC 1557

QY 520 LeuIleProProAspAlaIleProArgGlyIleTyrGluIleTyrLeuThrIleHis 539
Db 1558 CTCATACCCCGGAGTGCATCCCGGAGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

QY 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db 1618 AAGCCAGAGAGCTGAGCTTGCCTCTAGCTGCTGCAGACCTGCTGAGTCCAGTCGTT 1677

QY 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db 1678 AGCTGTGGGCCCCCAGGAGTCTGCTACCCCGGCGAGTCATCTTGCATGGACCACCTGT 1737

QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db 1738 GGAGAGCCCAAGCCCTGACAGCTGAGTCTGCGCCTCAAAAAGCAGTCTCGGAGGCGAGT 1797

QY 600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
Db 1798 TGGGAG---GATGTGCTGCACCTTGTGAGGAGTACCTTCCCACTCTACTACTGCGAG 1854

QY 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db 1855 CTGGAGCCCGGGGCTGCTATGCTTTCAGGAGCAGCTGGGCGCTTTCCTGCTGAGGA 1914

QY 640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla 659
Db 1915 GAGGCCCTCAGCGGCTGCGCACCAAGCGCTCAGGCTCTCTGTTGCTCCGCTGCC 1974

QY 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db 1975 TGTACGTCCCTTGAGTACACATCCGAGTGTACTGCTTACAGCACCCAGGCTCTC 2034

QY 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlnLeuLeuGlnGluProArgVal 699
Db 2035 AAGGAGGTGTGAGCTGGAGAGCAGCTAGGTGAGCAGCTGATCCAGGAGCTCGGCTC 2094

QY 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db 2095 CTCGACTTCAAAAGACAGTTACCAACCTACGCTCTCTCCATCCACGACGTGCCAGCTCC 2154

QY 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db 2155 CTGTGGAAGACCAAGCTACTTGTGCTACGACTACGAGATCCCTTTTACCACTCTGGAAC 2214

QY 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db 2215 GGCACCCAGCAGTATCTGCATCTGCACCTTCCCTTGGAGCGCATCAACGCCAGCACAGC 2274

QY 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db 2275 GACCTGGCTGCAAGGTGTGGGTGTGGCAGTGGAGGAGATGGGAGAGCTTCAACATC 2334

QY 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db 2335 AACTTCAACATCACTAAGACACAAAGTTTGTGAATTTGTGCTCTGGAGAGTGAAGGG 2394

QY 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db 2395 GGGGTCCCAAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAAG 2454

QY 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db 2455 ATCATCGCCAGTCTGACCCACCTCAGCGGGGCGGCGGAGTCTAGGCCAG 2514

QY 840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db 2515 AACTTCACTGACAGCCTTACTGCTCTTTGCTTCAAGCCCGCCCTCAGCCATG 2574

QY 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879

```

```

Db 2575 ATCTCAACCTATGGAGCGCGGCACTTCCCAACGCAACCTCGGCGAGCTGCAGCA 2634

QY 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 2635 GCTGTGGCGGACTGGGCAACCAACAGATGCTGCCCTTCTTACGGTGTGGAGCCGAGTGT 2694

RESULT 6
RNU87305 2697 bp mRNA linear ROD 15-MAY-1997
LOCUS Rattus norvegicus transmembrane receptor Unc5H1 mRNA, complete cds.
DEFINITION U87305
ACCESSION U87305
VERSION U87305.1 GI:2055391
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 2697)
AUTHORS Leonardo, E.D., Hinck, L., Masu, M., Keino-Masu, K., Ackerman, S.L. and
Tessier-Lavigne, M.
TITLE Vertebrate homologues of C. elegans UNC-5 are candidate netrin
receptors
JOURNAL Nature 386 (6627), 833-838 (1997)
MEDLINE 97271897
PUBMED 9126742
REFERENCE 2 (bases 1 to 2697)
AUTHORS Leonardo, E.D., Hinck, L., Masu, M., Keino-Masu, K. and
Tessier-Lavigne, M.
TITLE Direct Submission
JOURNAL Submitted (28-JAN-1997) Anatomy, UCSF, 513 Parnassus, San
Francisco, CA 94143-0452, USA
FEATURES
source
1. 2697
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="brain and ventral spinal cord"
/dev_stage="18 day embryo and 13 day embryo"
1. 2697
/codon_start=1
/product="transmembrane receptor UNC5H1"
/protein_id="AAB57678.1"
/db_xref="GI:2055392"
/translation="MAVRPGLWPVLGIIVLAALRGSGAQSATVANPVGANDLLP
HFLVEDVIVKPKVLLVCKAVPATQIFKCNQEMVQVDHVIERSIDSSGLPTM
EVRINVSROOVKVEGLEVMQCVAVSSGTTKSKAYIRIAYLRNFQEPFLAKEV
SLEQGLVLCRPPGPIPAEVENLRNEDLVDPDENVTITREHSLVVRQARLDATAN
YTCVAKNI VARRRSTSAAVIVYVNGGWSWTWMSVCSACGRGWQKRSCTNPAPLN
GGAFCEGQNVQKATATLCPVDGSSWSKSGACGLDCTHRSRECSDDPAPNGGEC
RGADLDTRNCTSDLCIHTASCPEDEVALYIGLVAVAVCLFLLALLGLIYCRKEGLDS
DVADSLILTSFGQPVSIKPSKADNPHTLIOPDLSTTTTYYOGLSCSRQDGPSPQL
SMHGLLSPLGSGRHTLHSSPTSEAEDFVSLSTQNYFSLPRGTSMAYGTFNPLGG
RLMIPNTGSLILPDAIRGKIYEILTHKPEDVRPLAGCTLLSPVSGCPGV
LUTRVLAMDHCGPSPDSWLRLKQCEGWEDVLHGEESPHESHLYCYOLEAGC
VYFTQLGRFALVGEALSVAAATKRLFLFAPVACTSLRYNIRVYCLHDTALKEAV
QLEKQGLGLIQEPRVLHFQDSYHNLRLSIHDVPSLSKSLLSYQEIPEVHIWNGT
QOYLCTFLERINASTDLACKVMWVGEGDQSFNFNITKTDRFAELALESEG
GYPALVGPSAFKIPFLIRQKIISLDPPCSRGRADWRTLAQKHLHLSHLSFFASKPSPT
AMILNWEARHPFNGLQGLAAVAGLQPDAGLFTVSEAC"

```

Alignment Scores:

Pred. No.:	1,28e-279	Length:	2697
Score:	4545.50	Matches:	852
Percent Similarity:	96.78%	Conservative:	19
Best Local Similarity:	94.67%	Mismatches:	26
Query Match:	94.96%	Indels:	43
DB:	10	Gaps:	3

US-09-970-944-2 (1-899) x RNU87305 (1-2697)

QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyLeValLeuAlaAlaTrpLeu 20
DB 1 ATGGCCGTCGGCCCGCGCTGTGGCCAGTCTCTCTGGCCATAGTCTCTGGCCGCTGCTT 60
QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
DB 61 CGTGGTTGGGGTCCCGACAGAGAGTCCACGGTGGCCATCCAGTCCCGTGGCCACCC 120
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
DB 121 GACCTGCTGCCACCTTCCTGGTAGAGCTGTAGAGCGTGTACATTTGTCAAGAAACAAGCCG 180
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
DB 181 GTGTGTGTGTGTGCAAGGCTGTGCTGCCACCCAGATCTTCTCAAGTGCATATGGGAA 240
QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
DB 241 TGGGTCCCGCAGGTGCGATCAGTAATTGAACGCAGCACCCAGCAGCAGCGGATTGCCA 300
QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
DB 301 ACCATGAGGTCCGTATCAACGTATCAGGCAGCAGCGATGAGAAAGTGTGGGCTGGAG 360
QY 121 GluTyrTrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAla 140
DB 361 GAATACTGGTGCCAGTGTGGCATGAGCTCTCGGGTACCACAAAGTCAAGAGGCC 420
QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
DB 421 TACATCCGGATTGCTATTTCGCGAAGACTTTGAGCAGGAGCCACTGGCCAAAGAGTG 480
QY 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
DB 481 TCACTGGAGCAAGCATTTGTACTACCTTGTGCCCCCCCCAGAAAGAAATCCCCCAGCTGAG 540
QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
DB 541 GTGAGGTGGCTTCGAAATAGAGACTCTGTGACCCCTCCCTCGATCCCAATGTGTACATC 600
QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
DB 601 ACGCGGAGACACGCTAGTCTGGTTCAGCCCGCTGGCCGACACGCGCCAACTACACC 660
QY 221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
DB 661 TGTGTGCCAAGAACATCGTAGCCGTGCGCGAAGCACTCTGTGACGCGGTCAATTTTAT 720
QY 241 ValAsnGlyGlyTrpSerThrTrpThrTrpSerValCysSerAlaSerCysGlyArg 260
DB 721 GTGAACGTTGGTGGTTCAGCTGAGTCTGAGTGGTCCGCTGTGAGCCCGCTGTGGGCGT 780
QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
DB 781 GCCTGGCAGAAACGAGCCGAGCTGACCAACCGGCACCTCTCAACGGGGCGGCCCTTC 840
QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
DB 841 TGTGAGGGCAGAAATGTC---CAGAAACAGCCTGCGCCACTCTGTGCGCAGTGTGATGG 897
QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
DB 898 AGCTGGAGTTCGTGGAGTAAGTGTGACCTGTGGGCTGTGATGACCCACCTGGCGGAGC 957
QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
DB 958 CGCGAGTCTCTGACCCAGCACCCCGCAATGGAGTGTGAGGAGTGTGGGCTGTGACCTG 1017
QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerClyProGluAspVal 360
DB 1018 GACACCCGCACTGTACCACTGCTGCTGCGACCGCTTCTGTGCCCGGAGGACGTG 1077
QY 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380

DB 1078 GCTCTCTACATCGCCTTGTGCTGTGGCTGTGCTCTTCTTGTGTGTGTGGCCCTT 1137
QY 381 IleLeuValTyrCysArgLysLysGlyLeuAspSerAspValAlaAspSerSerIle 400
DB 1138 GGACTCATTTTACTGTGCAAGAGAGGCTGAGCTCCGATGTGGCCGACTCGTCCATC 1197
QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
DB 1198 CTCACTCTGGGCTTTCAGCTGTGAGCATCAAGCCAGCAAGAGCAGACACCCCACTG 1257
QY 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
DB 1258 CTCACCATCCAGCAGACCTTCAGCACCCACCTACCACTTACCAGGCGACTCTATGTTCG 1317
QY 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
DB 1318 AGCAGGATGGACCCAGCCCAAGTTCAGCTCTCTAATGTCACCTGCTCAGCCCACTG 1377
QY 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluPheVal 479
DB 1378 GGGAGTGGCCGCTATAGTTGCACACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437
QY 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
DB 1438 TCCGCGCTCTCCACCCCAAACTACTTTCGTTCCCTGCGCGCGCACCCAGCAACATGGCC 1497
QY 500 TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
DB 1498 TACGGGACCTTCAACTCTCTCGGGGCGGCTGATGATCCCTAATACGGGATCAGGCTC 1557
QY 520 LeuIleProProAspAlaIleProArgLysIleTyrGluIleTyrLeuThrLeuHis 539
DB 1558 CTATACCCCGGATGCCATCCCGAGAGAAAGATCTACGAGATCTACCTCACACTGCAC 1617
QY 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
DB 1618 AAGCCAGAGACGTGAGTGTGCCCTAGTGGCTGTGAGACCCCTGTGAGTCCAGTCCGTT 1677
QY 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
DB 1678 AGCTGTGGGCCCCAGAGTCTCTCACC CGGCGAGTATCTTCCTGATGATGAGCACCTGT 1737
QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
DB 1738 GGAGAGCCAGCCCTGACAGCTGGAGTCTGGCTCAAAAAGCAGTCTCTGCGAGGCGAGT 1797
QY 600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
DB 1798 TGGGAG--GATGTGCTGACCTTGGTGAGGAGTCACTTCCACCTCTACTACTGCCAG 1854
QY 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
DB 1855 CTGAGGCGGGGCTGTATGTCTTACGAGAGAGCTGGGCGCGCTTTCCTCTGGTAGGA 1914
QY 640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla 659
DB 1915 GAGGCCCTCAGCTGGCTGCCACCAAGCGCTCAGGCTCTCTTGTGTTGCTCCGCTGCC 1974
QY 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
DB 1975 TGTAGTCCCTTGTAGTACAACATCCGAGTGTACTGCTTACACGACCCACCCAGCCTCTC 2034
QY 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlnLeuIleGlnGluProArgVal 699
DB 2035 AAGGAGGTGGTGGCTGGAGAGCAGTGTGTGAGCAGCTGATCCAGAGGCTCGCGCTC 2094
QY 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
DB 2095 CTGCACTTCAAAGACAGTTACCAACCTACGTCTCTCCATCCACGAGCTGCCAGCTCC 2154
QY 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739

```

Db      2155 CTGTGGGAAGACCAAGCTACTTGTACGTACCAGGAGATCCCTTTTATCCACATCTGGAAAC 2214
Qy      740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db      2215 GGCACCCAGAGTATCTGCATCTGCACCTTACCTTGGAGCGCATCAACGCCAGCACCAGC 2274
Qy      760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerile 779
Db      2275 GACCTGGCTGCAAGGTGTGGGTGGCGAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334
Qy      780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db      2335 AACITCAACATCCTAAGGACACAAAGGTTTGTGTAATTTGTGGCTCTGGAGAGTGAAGG 2394
Qy      800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db      2395 GGGGTCCAGCCCTGTGGGCCCCAGTGCCTTCAAGATTCCTCTCAITTCGGGCAAAAG 2454
Qy      820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db      2455 ATCATCGCCAGCTCTGGAGCCCAACCTCGCAGCGGGGGCGGCGACTGGAGAACTCTAGCCAG 2514
Qy      840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db      2515 AAACCTTCACCTGGACAGCCATCTTAGCTTCTTTCCTCCAAAGCCCGAGCTACAGCCATG 2574
Qy      860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db      2575 ATCTCAACCTATGGAGGACGAGCACTTCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634
Qy      880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db      2635 GCTGTGCGCGAGTGGGCCAACCAAGATGCTGGCTCTTTCAGGTGTGGAGGCGGAGTGT 2694

RESULT 7
BD057524
LOCUS      BD057524                3014 bp    DNA        linear    PAT 27-AUG-2002
DEFINITION Netrin receptors.
ACCESSION  BD057524
VERSION    BD057524.1 GI:22603130
KEYWORDS  JP 2001505062-A/1.
SOURCE    synthetic construct
           synthetic construct
           artificial sequences.
REFERENCE  1 (bases 1 to 3014)
            Lavigne,M.T., Leonardo, D.E., Hinck, L., Masu, M. and Masu, K.K.
TITLE     Netrin receptors
AUTHORS   Netrin receptors
JOURNAL   THE REGENTS OF THE UNIV OF CALIFORNIA
COMMENT   EN JP 2001505062-A 1 17-APR-2001;
           PD JP 2001505062-A/1
           PF 17-APR-2001
           PR 19-FEB-1998 JP 1998536840
           PI MARC FESSIER LAVIGNE, DAVID E LEONARDO, LINDSAY HINCK, MASAYUKI
           PI MASU,
           PI KAZUKO KEINO MASU
           PC C07K14/00, C07K14/00, C07K17/00, C07H21/02, C07H21/04, G01N33/53 CC
           CC Topology: Linear;
           PH Key Location/Qualifiers.

FEATURES             source
           1..3014
           /organism="synthetic construct"
           /mol_type="genomic DNA"
           /db_xref="taxon:32630"

ORIGIN
Alignment Scores:
Pred. No.:          1,48e-279          Length:          3014
Score:              4545.50           Matches:          852
Percent Similarity: 96.78%           Conservative:
Best Local Similarity: 94.67%         Mismatches:       26
Query Match:        94.96%           Indels:           3

```

```

DB:
US-09-970-944-2 (1-899) x BD057524 (1-3014)
Gaps:
3
Qy      1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db      1 ATGGCCGTCCGCGCCGCGCTGTGCCAGTGTCTCTGGGCATAGTCTCTCGCCCTTGGCTT 60
Qy      21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db      61 CGTGGTTCGGGTGGCCAGCAGTGCACGTGGCCAAATCCAGTGCCTGGCCGCGGTCGCAACCCC 120
Qy      41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db      121 GACCTGTGCCCCCACTTCTCTGTAGAGCTGAGGACGTGTACATTGTCAAGTCAAGACGCG 180
Qy      61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db      181 GTGTGTGTGGTGTGCAAGGTGTGCTTGCACCCAGATCTTCTTCAAGTCAATGGGGAA 240
Qy      81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db      241 TGGGTCCGCCAGGTGTGATCAGTAAATTGAACGACGACCGACAGCAGCAGCGGATGGCA 300
Qy      101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db      301 ACCATGAGGTCCGTATCAACGTATCGAGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360
Qy      121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db      361 GAATACTGGTGGCCAGTGTGTGGCATGGAGTCTCTCGGTACCCACCAAAAGTCAAGAGGCC 420
Qy      141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db      421 TACATCCGATTCCTATTTTGGCAAGAACTTTGAGCAGCAGGACCTTGGCCAGGAAGTG 480
Qy      161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db      481 TCACTGGAGCAAGGCAATTGTACTACCTTGTGCGCCCCCAGAGGAATTCGCCCGAGCTGAG 540
Qy      181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db      541 GTGAGTGGCTTCGAAATGAGGACCTCTGTGGACCCCTCCCTCGATCCCAATGTGTATC 600
Qy      201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
Db      601 ACGCGGAGCACAGCCCTAGTCTGTGCGTCAAGCCCGCTGCGCCGACACGCGCAACTACACC 660
Qy      221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240
Db      661 TGTGTGGCCAAAGAACATCGTAGCCCGTGGCCGGAAGCACCTCTGACGCGGTCTTGTATTAT 720
Qy      241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
Db      721 GTGAACGTTGGTGGTTCGACGTGACCTGTAGTGTCTCGTCTGACGCGCCAGCTGTGGCGGT 780
Qy      261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
Db      781 GGCTGGCAGAAACGGAGCGGAGCTGCACCAACCGCGACCTCTCAACGGGGCGGCTTC 840
Qy      281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
Db      841 TGTGAGGGGCAGAAATGTC---CAGAAACAGCGCTGCGCCACTCTGTGCCAGTGGATGGG 897
Qy      301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db      898 AGCTGGAGTTCGTGGAGTAAGTGTGTGAGCTGTGGGCTTTGACTGCACCCACCTGGCGGAGC 957
Qy      321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db      958 CGCGAGTGTCTTGACCCAGCAGCCCGCAATGGAGGTGAGAGTGTGGGGGTGCTGACCTG 1017
Qy      341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360

```

1018 GACACCGCAACAGTACAGTACCTCTGCTGCACACCGCTCTCTGCCCCGAGAGCGTG 1077
 361 AlaLeuTyrValGlyLeuLeuAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
 1078 GCTCTCTACATCGCCCTTGCTGCTGCTGCTGCTGCTCTCTCTGCTGCTGCTGCTGCT 1137
 381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerValAlaAspSerSerIle 400
 1138 GGACTCATTTACTGTCGAAAGGAAGAGGCTGAGCTCCGATGCTGGCCGACTCGTCCATC 1197
 401 LeuThrSerGlyPheGlnProValSerIleLeuProSerLysAlaAspAsnProHisLeu 420
 1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAAGAGAGAGAGAGAGAGAGAG 1257
 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThrThrThr 439
 1258 CTCACCATCCAGCCAGACCTCAGCACCCACCTACCTACCTACCTACCTACCTACCTAC 1317
 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
 1318 AGCAGGATGACCCAGCCCAAGTTCAGCTCTCTAATGCTCACTACCTGCTCAGCCCACTG 1377
 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrThrSerGluAlaGluPheVal 479
 1378 GGGAGTGGCCGCCATACGTTGCACACAGCTTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437
 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
 1438 TCCGCGCTCTCCACCCAAACTACTTTCGTTCCCTGCGCCGCGCACCAATGGCC 1497
 500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
 1498 TACGGGACCTTCAACTCTCTCGGGGGCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557
 520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
 1558 CTCATACCCCGGATGCCATCCCCGGAGGAAGATCTACGAGATCTACCTCAGACTGCAC 1617
 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
 1618 AAGCCAGAGACGTCGAGTTCGCCCTAGCTGGCTGTGACAGCCCTGCTGAGTCCAGTCGT 1677
 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
 1678 AGCTGTGGCCCCCAGGAGTCTGCTCACCAGGCGAGTCATCCTTGCAATGGACCACTGT 1737
 580 GlyGluProSerProAspSerTyrSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
 1738 GGAGAGCCCGCCCTGACAGCTGGAGTCTGGCTCAGAAAGCAGTCTCGGAGGGCAGT 1797
 600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
 1798 TGGGAG---GATGTGCTGCACCTTGGTGAGGAGTCACTCCCACTCTACTACTGCGCAG 1854
 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
 1855 CTGGAGGCGCGGCGCTGCTATGCTTCACGAGAGAGCTGGGCGCGCTTTCCTCGGTAGGA 1914
 640 GluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
 1915 GAGGCGCTCAGGTGGCTGCCACCAAGCGCTCAGGCTCTCTGTTGCTCCCGTGGCC 1974
 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
 1975 TGTACCTCCCTGAGTACAACTCCGAGTGTCTGCTACAGCAGACCCAGCAGCCTCTC 2034
 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuLeuGlnProArgVal 699
 2035 AAGGAGGTGGTGCAGCTGGAGAGAGAGCTAGGTGGAGAGCTATCCAGAGGCTCGCGTC 2094
 700 LeuHisPheLysAspSerTyrHisAsnLeuLeuArgLeuSerIleHisAspValProSerSer 719

2095 CTGCACCTCAAGACAGAGTTACCACAACCTACGCTCTCTCCATCCAGCAGCTGCCAGCTCC 2154
 720 LeuTyrLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTyrAsn 739
 2155 CTGTGAAGAGCAAGTACTTGTACGTACAGAGATCCCTTTTACCACTCTGGAC 2214
 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
 2215 GGCACCCAGCAGTATCTGCACCTTCACTCCCTGGAGCGCATCAAGCCAGCAGCAGCAGC 2274
 760 AspLeuAlaCysLysLeuTyrValTyrGlnValGluGlyAspGlyGlnSerPheSerIle 779
 2275 GACCTGGCTGCAAGGTGTGGGTGTGGAGTGGAGGAGATGGGAGAGCTTCAACATC 2334
 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
 2335 AACCTCAACATCATAGGACACAAGGTTTGCTGAATTGTGGCTCTGGAGAGTGAAGGG 2394
 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
 2395 GGGGTCCAGCCCTGGTGGGCCCGCAGTGCCTTCAAGATCCCTTCCTCATTTGGCAAAAG 2454
 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTyrArgThrLeuAlaGln 839
 2455 ATCATCGCAGTCTGGACCCACCTGCAGCGGGCGCGGCGGAGTCTAGGAGTCTAGCCAG 2514
 840 LysLeuHisIleAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
 2515 AAACCTTCACTGGACAGCCATCTTAGCTTCTTTCCTCCAGCCCGCCCTACAGCCATG 2574
 860 IleLeuAsnLeuTyrGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
 2575 ATCTCAACCTATGGGAGGAGCAGGCACTTCCCCAACGCGCACTCGGCCAGCTGGCAGCA 2634
 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
 2635 GCTGTGGCGGAGCTGGGCCAACAGATGCTGGCCTCTTACGGTGTGGAGGCGCGAGTGT 2694

RESULT 8
 AX367094 LOCUS AX367094 3580 bp DNA linear PAT 16-FEB-2002
 DEFINITION Sequence 13 from Patent WO0198354.
 ACCESSION AX367094
 VERSION AX367094.1 GI:18855296
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 Griffin,J.A., Kallick,D.A., Tribouley,C.M., Yue,H., Nguyen,D.B.,
 Tang,Y.T., Lal,P., Policky,J.L., Azimzai,Y., Lu,D.A., Graul,R.,
 Yao,M.G., Burford,N., Hafalia,A.J., Baughn,M.R., Bandman,O.,
 Patterson,C., Yang,J., Xu,Y., Warren,B.A., Ding,L. and
 Sanjanwalla,M.S.
 Receptors
 Patent: WO 0198354-A 13 27-DEC-2001;
 Incyte Genomics, Inc. (US)
 FEATURES
 source
 1. 3580
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 6052371CB1"

ORIGIN
 Alignment Scores: 6.64e-269 Length: 3580
 Pred. No.: 4379.50 Matches: 838
 Score: 93.22% Conservatave: 1
 Percent Similarity: 93.11% Mismatches: 2
 Best Local Similarity: 91.49% Indels: 59
 Query Match: 6 Gaps: 3
 DB:

US-09-970-944-2 (1-899) x AX367094 (1-3580)

Qy	1	MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu	20
Db	4	ATGSCCGTCGCGCGCGCTGTGGCAGCGCTCTTGGCATAGTCTCGCCGCTTGGCTC	63
Qy	21	ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro	40
Db	64	CGCGCTCGGTGCCCAGAGAGTCCACCGTGGCCACCCAGTGCCTGTGTGCCAACCGG	123
Qy	41	AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro	60
Db	124	GACCTGCTTCCCACTTCTCGTGTGAGCCGAGGATGTGTACATCGTCAAGAACCAAGCCA	183
Qy	61	ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu	80
Db	184	GTGCTGCTGTGTGCAAGGCGGTGCCGCCACGCAGATCTTCTCAAGTGCACCGGAG	243
Qy	81	TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro	100
Db	244	TGGTGGCCAGGTGGACCATGTATCGAGCGCACAGCGGAGCAGTGGGTGGCC	303
Qy	101	ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu	120
Db	304	ACCATGAGGTCCGCAATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTCGGGTGGAG	363
Qy	121	GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla	140
Db	364	GAATACTGGTCCCACTGCTGGCTGGCATGGAGCTCTCGGGCACCCACAGAGTCAGAAGGCC	423
Qy	141	TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal	160
Db	424	TACATCCGATAGCTATTGGCGAAGAACTTCGAGCAGGAGCCGCTGCCAAGAGGTG	483
Qy	161	SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu	180
Db	484	TCCCTGGAGCAGGCGATCGTCTGCCCTCGCTCCACCGAGGGCATCCCTCCACCCGAG	543
Qy	181	ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle	200
Db	544	GTGGAGTGGCTCCGGAACGAGGACCTGTGTGACCCGCTCCCTGGACCCCAATGTATACATC	603
Qy	201	ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr	220
Db	604	ACCGGGAGCACAGCTGGTGGTCGCAGAGCCCGCCTTGCTGTACACGGCCAACTACACC	663
Qy	221	CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr	240
Db	664	TGCGTGCCCAAGAACATCGTGGCACGTCGCCGCAGCGCTCCGCTGCTGTCTCATCGTCTAC	723
Qy	241	ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg	260
Db	723	-----	723
Qy	261	GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe	280
Db	723	-----	723
Qy	281	CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly	300
Db	724	-----GTGACGGC	732
Qy	301	SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer	320
Db	733	AGCTGAGCCCGTGGACAAGTGTGCGCTGTGGCTGGACTGCACCCACTGGCGGAGC	792
Qy	321	ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu	340
Db	793	CGTGAGTGCTCTGACCCAGACCCCGCAACGAGGGGAGGAGTCCAGGGCACTGACCTG	852
Qy	341	AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal	360

Db	853	GACACCCGAACTGTACAGTGACCTCTGTGTACACTGCTCTGTGGCCCTGAGACGTG	913
Qy	361	AlaLeuTyrValGlyLeuIleAlaValCysLeuValLeuLeuLeuValLeu	380
Db	913	GCCCTCTATGTGGGCTCATGGCGTGGCGTCTGCCTGGTCTCTGCTGTGCTTGTGCTC	972
Qy	381	IleLeuValTyrCysArgGlyLeuGlyLeuAspSerAspValAlaAspSerIle	400
Db	973	ATCCTCGTTTATGCGGAAGAGGAGGGGTGGACTCAGATGTGGTGA	1033
Qy	401	LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAspProHisLeu	420
Db	1033	CTCACCTCAGGCTTCCAGCCCGTCAAGCATCAAGCCCAAGAGCAGACAACCCCATCTG	1099
Qy	421	LeuThrIleGlnProAspLeuSer--ThrThrThrThrTyrGlnGlySerLeuCysPro	439
Db	1093	CTCACCTCAGCGGGACCTCAGACCAACCAACCACTACCAAGGGCAGTCTCTGTGCC	1155
Qy	440	ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu	459
Db	1153	CGGAGGATGGGCCAGGCCCAAGTTCAGCTCAACCAATGGCACCTGCTCAGCCCCCTG	1212
Qy	460	GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal	479
Db	1213	GGTGGCGGCGCCACACACTGCACACAGACTCTCCACCTCTGAGGCGAGAGTTCGTC	1272
Qy	480	SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr	499
Db	1273	TCCGCGCTCTCACCCAGAACTACTTCGCTCCTGCCCCGAGGACCAACATGACC	1332
Qy	500	TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu	519
Db	1333	TATGGGACCTTCAACTTCTCGGGGCGCGGTGATGATCCCTATACAGGAATCAGCGTC	1392
Qy	520	LeuIleProProAspAlaIleProArgGlyIysIleTyrGluIleTyrLeuThrLeuHis	539
Db	1393	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAGCGTCAC	1452
Qy	540	LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal	559
Db	1453	AAGCCGAGACGTGAGTTGCCCCCTAGCTGGCTGTGAGACCCCTGCTGAGTCCCATCGTT	1512
Qy	560	SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys	579
Db	1513	AGCTGTGGACCCCTGCGCTCTGCTCACCGGGCCACTCATCTCTGGCTATGGACCATGT	1572
Qy	580	GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer	599
Db	1573	GGGAGAGCCAGCCCTGACAGCTGGAGCGCTCGCCCTCAAAAAGCAGTCTGTCAGGGCAGC	1632
Qy	600	TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln	619
Db	1633	TGGAG--GATGTGCTGCACCTGGGCGAGAGGGCGCCCTCCACCTCTACTCTGCCAG	1689
Qy	620	LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly	639
Db	1690	CTGAGGCCAGTGTGCTACGCTTTCACCGAGCAGCTGGCGCGCTTTCGCTTGGTGGGA	1749
Qy	640	GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla	659
Db	1750	GAGGCCCTCAGCGTGGCTGCGGCAACGCGCTCAAGCTGTCTCTGTTGCGCGGTGGCC	1809
Qy	660	CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu	679
Db	1810	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCTGTGATGACACCCACGATGCATC	1869
Qy	680	LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal	699
Db	1870	AAGGAGTGGTGGCTGGAGACAGCTGGGGGGACGCTGATCCAGGAGCACCGGTC	1929
Qy	700	LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer	719
Db	1930	CTGCATCTCAAGACAGTATACCAACCTCGGCTATCATTCACCATGTGCCAGCTCC	1989

QY	720	LeuTriplysSerLysLeuValSerTyrGlnGluLeuProPheTyrHisIleTrrPAsn	739
Db	1990	CTGTGGAAGAGTAAGTCCCTTGTTCAGCTACGAGGATCCCTTTATCATCATCTGGAAT	2049
QY	740	GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer	759
Db	2050	GGCACGCGAGGGTACTTGCACCTTCACCTTGAGGCGTGTACCCCGAGCAGCTAGT	2109
QY	760	AspLeuAlaCysLysLeuTriplValTrrPglNValGluGlyAspGlyGlnSerPheSerIle	779
Db	2110	GACCTGCGCTGCAAGCTGTGGTGTGCGAGTGCGAGCGGCGAGGCTTCAGCATC	2169
QY	780	AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuAlaLeuGluSerGluAla	799
Db	2170	AACTTCAACATCAACCAAGGACACAAAGTGTTCGCTGAGCTGTGGCTCTGGAGAGTGAAGCG	2229
QY	800	GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleAtrGlnLys	819
Db	2230	GGGGTCCAGCCCTGTGTGGCCCGAGTGCCTTCAAGATCCCTTCCTCATTCGCGCAAG	2289
QY	820	IleIleSerSerLeuAspProProCysArgArgGlyAlaAspThrArgThrLeuAlaGln	839
Db	2290	ATAATTTCCAGCTGGACCCACCCCTGTAGGCGGGTGCAGACTGGCGACTCTGGCCAG	2349
QY	840	LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet	859
Db	2350	AAATCCACCTGGACGACCATCTCAGCTTCTTTGCTCCAGCCGAGCCCGACAGCCATG	2409
QY	860	IleLeuAsnLeuTrrGluAlaAtrGHisPheProAsnGlyAenLeuSerGlnLeuAlaAla	879
Db	2410	ATCCTCAACCTGTGGAGCGCGGACCTTCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2469
QY	880	AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerLgrluAlaGluCys	899
Db	2470	GCACTGGCTGGAGTGGCCAGCCAGAGCTGGCTTTCACAGTGTGGAGGCTGAGTGC	2529
RESULT 9			
BC058084		3844 bp mRNA linear ROD 21-OCT-2003	
LOCUS		Mus musculus unc-5 homolog A (C. elegans), mRNA (cdna clone	
DEFINITION		MGC:66671 IMAGE:6813463), complete cds.	
ACCESSION		BC058084	
VERSION		BC058084.1 GI:34784158	
KEYWORDS		MGC.	
SOURCE		Mus musculus (house mouse)	
ORGANISM		Mus musculus	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
		1 (bases 1 to 3844)	
		Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,	
		Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,	
		Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,	
		Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,	
		Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,	
		Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,	
		Scheetz,T.E., Brownstein,M.J., Ustin,I.B., Toshiyuki,S.,	
		Carninci,P., Prange,C., Raja,S.S., Loquellano,N.A., Peters,G.J.,	
		Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,	
		McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,	
		Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,	
		Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,	
		Fahy,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,	
		Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,	
		Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,	
		Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,	
		Butterfield,v.S., Krzywinski,M.I., Skalska,U., Smalus,D.E.,	
		Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.	
		Generation and initial analysis of more than 15,000 full-length	
		human and mouse cdna sequences	
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
MEDLINE		22388257	
PUBMED		12477932	

REFERENCE		2 (bases 1 to 3844)
AUTHORS		Strausberg,R.
TITLE		Direct Submission
JOURNAL		Submitted (08-SEP-2003) National Institutes of Health, Mammalian
		Gene Collection (MGC), Cancer Genomics Office, National Cancer
		Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
		USA
REMARK		NH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT		Contact: MGC help desk
		Email: cgapsb-remail.nih.gov
		Tissue Procurement: Dr. Jim Lin, University of Iowa
		cDNA Library Preparation: M. Bento Soares, University of Iowa
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
		DNA Sequencing by: Genome Sequence Centre,
		BC Cancer Agency, Vancouver, BC, Canada
		info@cgsc.bc.ca
		Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
		Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
		Letitia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
		Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
		Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
		Schein, Duane Smalus, Michael Smith, Lorraine Spence, Jeff Stott,
		Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
		George Yang, Scott Zuyderduyn, Marco Marra.
		Clone distribution: MGC clone distribution information can be found
		through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
		Series: IRAC Plate: 126 Row: b Column: 11
		This clone was selected for full length sequencing because it
		passed the following selection criteria: matched mRNA gi: 23346570.
FEATURES		Location/Qualifiers
source		1..3844
		/organism="Mus musculus"
		/mol_type="mRNA"
		/strain="C57BL/6"
		/db_xref="taxon:10090"
		/clone="MGC:66671 IMAGE:6813463"
		/tissue_type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc"
		/clone_lib="NIH_BMAP_FWO"
		/lab_host="DH10B"
		/note="Vector: pYX-ASC"
gene		1..3844
		/gene="Unc5a"
		/db_xref="LocusID:107448"
		/db_xref="MGI:894682"
		252..2780
		/codon_start=1
		/product="Unc5a protein"
		/protein_id="AAH58084.1"
		/db_xref="GI:34784159"
		/db_xref="LocusID:107448"
		/translation="MAVRPGIWPALLGIVLTAWLRGSGAQSATVANPVGANPDLLP
		HLVEPEDVIVKPKVLLCKAVPATQIFKCGEWRVQDVHVERSTGSSGLPTM
		EFLINVRQKQVKFGLLEYWCQCVAWSSGTTKSKAYIRIAYLRKNFOEPLAKEV
		STGCIVLPCRPPEGIIPAEVWLRNEDLPDLPNVYITREHSLVVRQALADTN
		YLVCAKIVARRRSASAVIVVDGSPWSPKSCAGCLDCTHMRSCSPAPRNGGE
		ECRADLDTNCTSDCLHTSSGPEVALYGLVAVAVCLILLVLLVLYCRKKEGL
		DSDVADSILTSQGPVSKPSKADNPHELLIQDLSLTITTYGSLCPCPDGSPKPE
		QLSNGHLLSPGSRGTHHSPTSEADDFVSRUSTQNYFRSLPRGTSNMAIGHNFI
		CGRLMINTGILIPDAPRGKIYIYTLTKHPEDVRLPLAGCOTLLSPVSCGPP
		GVLITRPVILAMDHGSPDSWSLRKLKQSCGSEWDLHGLSESPSHLYCQLEAG
		ACVYEQLGOLIOEPRLHFKDSYHRLRLSIHDVPSSLKMSKLVSQEIPIFYHIW
		VVOLEKGLGOLIOEPRLHFKDSYHRLRLSIHDVPSSLKMSKLVSQEIPIFYHIW
		GTQOYLCTFLERNASTSDLACKVWVQEGQGFNINFINITKDFRAEMIALES
		EGVPALVGVSAFKPIPLIRKQITLSDPPCSGADWRTLAQKLHSLHSEFFASKPS
		PTAMILNWFARHPFNIGLQAAAVALGQPDAGLFTVSEAC"
		714..926
		/note="IG; Region: Immunoglobulin"
		/db_xref="CDD:smart00409"
		984..1130
		/note="TSPI; Region: Thrombospondin type 1 repeats"
		/db_xref="CDD:smart00209"
		1566..1877
		misc_feature
		misc_feature
		misc_feature

/note="ZUS; Region: Domain present in ZO-1 and Unc5-like
netrin receptors"
/db_xref="CDD:smart00218"
2495..2684
misc_feature

/note="DEATH; Region: DEATH domain, found in proteins
involved in cell death (apoptosis). Alpha-helical domain
present in a variety of proteins with apoptotic functions.
Some (but not all) of these domains form homotypic and
heterotypic dimers"
/db_xref="CDD:smart00005"

ORIGIN

Alignment Scores:

Pred. No.: 1-93e-262 Length: 3844
Score: 4278.50 Matches: 809
Percent Similarity: 92.00% Conservative: 19
Best Local Similarity: 89.89% Mismatches: 13
Query Match: 89.38% Indels: 59
DB: 10 Gaps: 3

US-09-970-944-2 (1-899) x BC058084 (1-3844)

QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
DB 252 ARGCGCGTCCGCGCGCGCTGTGGCAGCGCTCTCTGGGCATAGTCTCACTGCGCTGGCTT 311
QY 21 ArgGlySerGlyAlaGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
DB 312 CGTGGTTCGGGTGCCACAGAGTCCACAGTGGCCAAACCAGTGCCTGGTGGCCAAACCG 371
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
DB 372 GACCTGTCTGCCCACTTCTGGTAGCGCGGAGGACGTGTACATGTCTCAAGAACAGGCC 431
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
DB 432 GTGCTGTGGTGTGCAAGGCTGTGCCGCCACCCAGATCTTCTCAAGTGCAACGGGAA 491
QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
DB 492 TGGGTTCGCCAGTGTGATCAGTCAATTGAACGACAGCTGACGGCAGCAGCGGATTCGCA 551
QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
DB 552 ACCATGAGGTCCGATCAACGTATCAAGGACAGCGTCCGAGAAGTGTGTTGGCGTGAG 611
QY 121 GluTyrTrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAla 140
DB 612 GAGTACTGGTGTCCAGTGTGTGGCATGGAGCTCTCTCAGGAACACCAAGAGCCAGAGGCC 671
QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
DB 672 TACATCCGATTCGCTATTTCGCCAAGAACCTTTGAGCAGGAGCGCTGGCCCAAGGAAGT 731
QY 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
DB 732 TCACTGAGCAAGGCAATTGTGCTACCTTGTGCCCGCCGAGAGAAATCCCGCCAGCTGAG 791
QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
DB 792 GTGAGTGGCTCCGAAATGAGGACCTCGTGAGACCCCTCCCTCGACCCCAATGTGTATC 851
QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
DB 852 ACAGGAGGACAGCTATGTCTGTGGCAGCGCCGCTGGCCGACACTCCCACTACACC 911
QY 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaValIleValTyr 240
DB 912 TGGGTGCCAAGAACATCGTGGCCCGTCCGAGAGCGCTCTCGCGCGCTCATTTGTTAT 971
QY 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
DB 971 ----- 971

QY 261 GlyTrpGlnLysArgSerArgSerCysThrAenProAlaProLeuAsnGlyGlyAlaPhe 280
DB 971 ----- 971
QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
DB 972 -----GTGGATGGG 980
QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
DB 981 AGCTGGAGCCCATGGAGTAAGTGTACGCTCGGGCTTGACTGCACCCACTGGCGGAGC 1040
QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
DB 1041 CGGGAGTGTCTCGACCCAGCGCCGCAACGAGAGTGAGGAGTGC CGGGGTCTGACCTG 1100
QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
DB 1101 GACACCCGCAACTGTACCAAGTACCTCTGCTGCACACCTCTTCGGGCCCGGAGACGTG 1160
QY 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
DB 1161 GCTCTCTACATCGGCTCGTCCGCGTGGCGGTGGCTTCATCTTCTGCTGCTGCTGCTC 1220
QY 381 IleLeuValTyrCysArgLysGlyGlyLeuAspSerAspValAlaAspSerSerIle 400
DB 1221 GTCTCTCATCTCTGCGCAAGAGAGAGACTGGACTCAGACGTGGCTGACTCATCCATC 1280
QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
DB 1281 CTTTACCTCAGCTTCCAGCCTGTGAGCATCAAGCCCAAGAGCAGACAATCCCCATCTG 1340
QY 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
DB 1341 CTCACCATCCAAACCGGACCTCAGCACCAACAGCCACCTACAGGGGAGGCTGTGTGCC 1400
QY 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
DB 1401 CGGCAGGATGAGACCCAGCCCAAGTTCAGCTCTCTAATGTGTACCTGTCTCAGCCCATG 1460
QY 460 GlyGlyArgHisThrLeuHisSerSerProThrSerGluAlaGluGluPheVal 479
DB 1461 GGCAGTGGCGGCCATACGCTGCACACAGCTCCCGCCACCTCTGAGGCTGAGGACTTCGTC 1520
QY 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
DB 1521 TCCCGCTCTCCACCCCAAACTACTTTCGTTCTCTGCCCGCGGTACAGCAACATGCC 1580
QY 500 TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
DB 1581 TATGGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAACACAGGAATCAGCCTC 1640
QY 520 LeuIleProProAspAlaIleProArgGlyValIleTyrGluIleTyrLeuThrLeuHis 539
DB 1641 CTATATCCCGCGAGCGCATCCCGCGAGGAAGATCTACAGAGATCTACCTCACTCTGCAC 1700
QY 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
DB 1701 AAGCCAGAAAGCTGAGTTGCCCTAGCTGGCTGTGACACCTCTGAGTCTCTATCGTT 1760
QY 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
DB 1761 AGCTGTGGGCGCCAGGAGTCTGTCTCACCGGCCAGTCACTCTGCCATGGACCACTGC 1820
QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
DB 1821 GGGGAGCCAGTCCCGACAGCTGAGCCTCGCCTCAAAAAGCAGCTCTGTGAGGCGAGC 1880
QY 600 TrpGluAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGln 619
DB 1881 TGGGAG---GACGTGTGCACCTTGTGTGAGGAGTGCCTCTCATCTACTACTGCCAG 1937

Qy 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
 Db 1938 CTGAGGCGGGGCTGCTATGCTTTACCGAGCAGTAGGCGCTTTGGCCCTGGTGGGA 1997
 Qy 640 GluAlaLeuSerValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 659
 Db 1998 GAGGCCCTCAGCTGGCTGGCCACCAAGCGCTCAGGCTCTCTTGTGGCCCTGTGGCC 2057
 Qy 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
 Db 2058 TGTACGTCCTCGAGTACCAACATCCGAGTGTACTGCTGCAGCAGCACCCAGATGCTCTC 2117
 Qy 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
 Db 2118 AAGGAGGTGGTGGAGTGGAGAGCAGCTGGGTGGACAGCTGATCCAGGAGCCCGCTGTC 2177
 Qy 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
 Db 2178 CTGCACCTCAAGACAGTATACCAACCTAGCTGTCTGCATCCAGCAGCTGCCAGCTCC 2237
 Qy 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
 Db 2238 CTGTGGAAGAGCAGCTCTGTGCTAGCTACCAAGAGATCCCTTTTACCATCTGGAAT 2297
 Qy 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
 Db 2298 GGCACCTACGAGTATCTGCTGACCTTACCTCCCTGGAGCGCTCAATGCCAGCACCCAGC 2357
 Qy 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
 Db 2358 GACCTGGCTGCAAGGTGGGTGGGAGGTGGAGGTGGAGGTGGAGGTGGAGGTGGAGGT 2417
 Qy 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
 Db 2418 AACTTTAATCCTCAAGACACGAGGTGTGCTGAATGCTGGCTCTGGAGAGTGAAGG 2477
 Qy 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
 Db 2478 GGGGTCCAGGCGCTGGTGGGCGCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAAG 2537
 Qy 820 IleLeuSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
 Db 2538 ATCATACCGCTGACCCACCTTCAGCGCGGGCGCGGACTGCGGAACCTTAGCCAG 2597
 Qy 840 LysLeuHisLeuAspSerHisLeuSerPheAlaSerLysProSerProThrAlaMet 859
 Db 2598 AAACCTTCACTGGACGAGCATCTTAGCTTCTTTGGCTCCCAAGCCCGCCCTACAGCATG 2657
 Qy 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaLa 879
 Db 2658 ATCCCTCAACCTATGGAGCGCGGACCTTCCCAAGCGCAACCTTGGCCAGCTGGCCGA 2717
 Qy 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
 Db 2718 GCTGTGGCGGAGTGGCCAGCAGCATGTGGCTCTTCCAGCTTCCAGGCGCGAGTGC 2777
 RESULT 10
 AK122575
 LOCUS Mus musculus mRNA 4294 bp mRNA linear ROD 15-MAR-2003
 DEFINITION Mus musculus mRNA for mKIAA1976 protein.
 ACCESSION AK122575
 VERSION AK122575.1 GI:28972881
 KEYWORDS FLI_CDNA.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Aizawa,H., Yuasa,S.,
 Nakajima,D., Nagase,T., Ohara,O. and Koga,H.
 Prediction of the coding sequences of mouse homologues of KIAA
 gene: II. The complete nucleotide sequences of 400 mouse
 KIAA-homologous cDNAs identified by screening of terminal sequences

of cDNA clones randomly sampled from size-fractionated libraries
 DNA Res. 10, 35-48 (2003)
 2 (bases 1 to 4294)
 AUTHORS Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
 DIRECT SUBMISSION
 TITLE Submitted (07-FEB-2003) Hisaaki Koga, Kazusa DNA Research
 Institute, Laboratory for Genome Informatics; 2-6-7
 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
 (E-mail:mouse@kazusa.or.jp, Tel.:81-438-52-3919, Fax:81-438-52-3918)
 The CREATE program supported by Japan science and technology
 corporation; cDNA full insert sequencing: Kazusa DNA Research
 Institute; cDNA library construction, clone selection and 5'- &
 3'-end one pass sequencing.
 FEATURES
 Location/Qualifiers
 1..4294
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="mbh03993"
 /tissue_type="brain"
 /dev_stage="adult"
 /note="vector:modified pBC SK+"
 1..4294
 /gene="mKIAA1976"
 /cDS_start=1
 /evidence="not experimental"
 /product="mKIAA1976 protein"
 /protein_id="BAC65857.1"
 /db_xref="GI:28972882"
 /translation="SLVTPALLKTTLSAASSGPDVALYGLVAVAVCLILLLVLV
 LIYCRKKEGLSDVADSSILTSFGQPSVSKSKADNHLIIQDLSITTTTYQGLC
 PRQPGSPKQLEKQLGQIEPRVLFHFKDSYHNLSIHEDVPSLWKSLLVSYQE
 IPFYHNGTQYLLHCTFLERVNASTDACKVWQVEGDSFNFINFKDTRF
 AEMALISEGVPALVGPFAFKIPFLIRQKITSLDPPCSRGAQWRTLAQKLDHSL
 SFFASKPSPTAMILNLWEARHPNGNLQQLAAAVAGLQPDAGLFTVSEAC"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,216-182 Length: 4294
 Score: 3024.50 Matches: 632
 Percent Similarity: 49.09% Conservative: 15
 Best Local Similarity: 47.95% Mismatches: 15
 Query Match: 63.18% Indels: 657
 DB: 10 Gaps: 4
 US-09-970-944-2 (1-899) x AK122575 (1-4294)
 Qy 2 AlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeuArg 21
 Db 2 GCGCTCCGCGCGCGCTGTGGCCAGCGCTCTGGGCGATAGTCTCTACTGCTGGCTTCT 61
 Qy 22 GlySerGlyAlaGlnSerAlaThrValAlaAlaAsnProValProGlyAlaAlaAsnProAsp 41
 Db 62 GGTTCGGGTGCCAGCAGAGTGCACAGTGCACACCCAGTGCCTGGTCCCAACCGGAC 121
 Qy 42 LeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysProVal 61
 Db 122 CTGCTGCCCACTCTCTGTGTAGCGGAGGAGCGTGTACATTTGTCAAGAACAGCCCGTG 181
 Qy 62 LeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrp 81
 Db 182 CTGCTGGTGTGAAGCTGTGCCCACTGTCCCACTGTCAAGTGCACCGGGATGG 241
 Qy 82 ValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluProThr 101
 Db 242 GTTCGCCAGTCCATCCATTCATTTGAACGACACTGACGCGGAGCAGCGATTCGCAAC 301
 Qy 102 MetGluValArgIleAsnValSerArgGlnValGluLysValPheGlyLeuGluGlu 121

Db 302 ATGGAGGTCGGATCAACGTTATCAAGGACGAGGTCTGAGAAAGTCTTTGGGCTGGAGGAG 361
 Qy 122 TyrTrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAlaTyr 141
 Db 362 TACTGGTGCAGTGTGTGGCATGGAGCTCTCTCAGGAACACCAAAAGCCAGAGGCGCTAC 421
 Qy 142 IleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSer 161
 Db 422 ATCCGGAATTCCTATTTCGCAAGAACTTTGAGCAGGAGCGCTGGCCCAAGGAAGTGCA 481
 Qy 162 LeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGluVal 181
 Db 482 CTGAGCAAGGCAATGTGTCTACCTTGTGCGCCCGCCGGAAGGAATCCCGCCAGCTGAGGTG 541
 Qy 182 GluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThr 201
 Db 542 GAGTGGCTCCGAATAGAGCACTGTGGACCCCTCCCTCGACCCCAATGTGTATCATACA 601
 Qy 202 ArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCys 221
 Db 602 CGGAGACACAGCTAGTGTGCGGACGCGCGCTGGCCGACACTGCCAACTACACCTGC 661
 Qy 222 ValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyrVal 241
 Db 662 GTGGCCCAAGAAACATCGTGGCCCGCTGCGGAAGCGCTCTGCGGCGCTCATTTGTTATGTG 721
 Qy 242 AsnGlyGlyTrpSerThrTrpThrTrpSerValCysSerAlaSerCysGlyArgGly 261
 Db 722 AACGGTGGGTGCAGTGTGACCGAGTGTGCTGTCGACGCCAGCTGTGGCGGTGGC 781
 Qy 262 TrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyValaPheCys 281
 Db 782 TGGCAGAAACGAGCGGAGCTGCACCAACCCGCGACCTCTCAACGCGGCGCTTCTGT 841
 Qy 282 GluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySer 301
 Db 842 GAGGGGCAGAAATGTC---CAGAAACACGCTGCTGTGCCAGTGGATGGGAGC 898
 Qy 302 TrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArg 321
 Db 899 TGGAGCCATGGAGTAGTGTGAGCTGCGGCTTGTGACGCCACTGGCGGAGCGG 958
 Qy 322 GluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAsp 341
 Db 959 GAGTGTCTCCGACCGCGCCGCAACGAGGTGAGGAGTGGCGGGTGTGCTGACTGGAC 1018
 Qy 342 ThrArgAsnCysThrSerAspLeuCysValHisSerAla 354
 Db 1019 ACCCGCAACTGTACCACTGACGTGACCTGTGCTGCACAGTGA-GTCTTTCTCTGCCCGGAGGCTC 1077
 Qy 354 ----- 354
 Db 1078 ACCTCGCTGTAGACCTATGGCTGTGGTACGGGACGCCACTCTCTCGCTGCCTGTGTCTC 1137
 Qy 354 ----- 354
 Db 1138 TCCAAAGCTTGGAGCCAGGGTCCCTTTCAATCTCTAGCACACTCTCTCCCTGTACACA 1197
 Qy 354 ----- 354
 Db 1198 GACACACCACTTCCCTCTGAGACACACAGAGATGTGAATCTTGACCCCTGTATACCAA 1257
 Qy 354 ----- 354
 Db 1258 CTGCTCCCCCGGGCTTTGAAATTATCCCGTCCCGCAAGCTCACATAAGTCTCTGTGGG 1317
 Qy 354 ----- 354
 Db 1318 GGTGCTGGGCTGTGAGGCGCTGTTAATTACAGGCCATAACTAACAAATTAAGATCTCTG 1377
 Qy 354 ----- 354
 Db 1378 TTGGATTTTTTTTTTCTTTAGACTACACCTCCCTGCTTCTTCAATTAAAAACCTAA 1437

Qy 354 ----- 354
 Db 1438 TTATCTGTGTCAATTCCTTCTGTATGGACATTTCTTCCATTATGCAAAACTCTTAAAAAT 1497
 Qy 354 ----- 354
 Db 1498 CGATAGTGTCTTCTCCAAAGCAGTTGGTCCCCCAACCCCAAGCTTTCAGGGCTGGTGA 1557
 Qy 354 ----- 354
 Db 1558 TACAAAGGGCTCACACTGTCTCTGGAACACCCCGAAGAACTGACCCCACTGATCTGCC 1617
 Qy 354 ----- 354
 Db 1618 TATTTTCTATTCTATAATGAAGTAGCAAAAGAAATTTATTAGTTTATAGAGCTCAAAAT 1677
 Qy 354 ----- 354
 Db 1678 CCAATAGCGTGTGCTGGCTCTGGCAAAATGTGTCTTCCACCCCAAGCCACATTTGCTTGG 1737
 Qy 354 ----- 354
 Db 1738 TGGCAGATAGTACCACATCAGTACTGTAGACAAGAGAGAGATCACGTCTCAGGAAGAGAG 1797
 Qy 354 ----- 354
 Db 1798 AAAGACACACAGAGACAGGTAGACACAGAAAGATGGACAGCCCCCATGGTCTAAGACA 1857
 Qy 354 ----- 354
 Db 1858 GCACATAATGATGTAAATGCCTCCCCACCTGGGGGCCCACTCTTAAAGGCGCCACTGCT 1917
 Qy 354 ----- 354
 Db 1918 GCTCATATTGCCAATGGGACACACCTGTCTCAATATGAATCCTTAGGAGGATGTGC 1977
 Qy 354 ----- 354
 Db 1978 AAACTATAGCACCATACTTAGGGTCTCTCAGGAGTTTCACTTGTGCCCGACACCCCTGACC 2037
 Qy 354 ----- 354
 Db 2038 CCCAGCCTAGGAGCCAAAGGATATACCACTCACCCCGGTGAAGTAGTCTTCCCTGTC 2097
 Qy 354 ----- 354
 Db 2098 TCCACTCTGCTGTGAGGGTGTGACTTGTCTCAATCCCAAGCAAGTTGTGTGGGTATAG 2157
 Qy 354 ----- 354
 Db 2158 AACTAGGATGTACCCACCAATACCAAGCTCCCTGCCACGTCGCCACTGCCCCAGGGGGTGC 2217
 Qy 354 ----- 354
 Db 2218 TGAGATCTTATCTCTGGGGATAAACCATGGGCTCTCTCTGAAGCCTCTTACATCCCT 2277
 Qy 355 -----SerGlyProGluAspValAlaLeu 362
 Db 2278 GCCATTCTCAAGACCACACTTTTCTCTCGCAGCGCTCTTCGGCCCCCGAGGACGTGGCTCTC 2337
 Qy 363 TyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeuLeuThr 382
 Db 2338 TACATCGGCTCTGTCGCCGTGTGCTCTCATCTTGTCTGTCTGTCTGTCTGTCTGTCT 2397
 Qy 383 ValTyrCysArgLysGlyLeuAspSerAspValAlaAspSerSerIleLeuThr 402
 Db 2398 ATCTACTCGCGAAGAAAGGAGACTGGACTCAGATGTGGCTGACTCATCCATCTTACC 2457
 Qy 403 SerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeuThr 422
 Db 2458 TCAGGCTTCCAGCTGTACAGCATCAAGCCAGCAAGAGAGACATCCCATCTGTCTCACC 2517

Web site: <http://www.nisc.nih.gov/>

Contact: nisc mgc@nigri.nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Young, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E.B. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 26 Row: g Column: 22.

FEATURES

source

Location/Qualifiers
 1..2688
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4126760"
 /tissue_type="Brain, neuroblastoma"
 /clone_lib="NTH MGC_19"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 <1..1627

CDS

/codon_start=2
 /product="Unknown (protein for IMAGE:4126760)"
 /protein_id="AAH09333.2"
 /db_xref="GI:40226528"
 /translation="DVALYGLIAVAVCLVLLLVLLVYCRKEGLSDVDSSILT
 SGFQVSIKSKADNPHLLTTPDLSTTTTQGSCLPRQDQSPKFLTNHGLISPL
 GGGTHLHSSPTSEAEFVSRSTQNYFRSLPRGTSNMTYTFNFGRLMIPNTGI
 SLLIPDAIPRGIYEIYLTHKPEDVRLPLAGCTLLSPVSCGPPGLLTRVILA
 MDHCSPSPDSRLRKQSCGSEWDLHGEAPSHLYCQLEAKCYFTSQLR
 FALVGEALSVAARLKLILPAPVACTSLEINIRVCLHDTHDHAKVQLEKQGGQ
 LIQPRVLFKDSVHNLFSLHDPSSLSKLIVSYOEIPFYHWNQYRQLHCTFT
 LERVSPSTDLACKLWQVEGQGSFINFNITKTRFAELLAEAGVPAUVGFS
 AFKIPFIQRIISLDPCCRRGADWRTLAQKLDHLSLFSFASKRSPMTAMINLWEA
 RHPFNGNLSQLAAVAGLQDPDAGLFTVSEAC"
 413..724
 /note="ZUS; Region: Domain present in ZO-1 and Unc5-like
 netrin receptors"
 /db_xref="CDD:smart00218"
 1343..1594
 /note="DEATH; Region: DEATH domain, found in proteins
 involved in cell death (apoptosis)"
 /db_xref="CDD:smart0005"

misc_feature

/note="ZUS; Region: Domain present in ZO-1 and Unc5-like
 netrin receptors"

misc_feature

/note="DEATH; Region: DEATH domain, found in proteins
 involved in cell death (apoptosis)"
 /db_xref="CDD:smart0005"

ORIGIN

Alignment Scores:

Pred. No.: 7,69e-170 Length: 2688
 Score: 2819.00 Matches: 540
 Percent Similarity: 99.63% Conservative: 0
 Best Local Similarity: 99.63% Mismatches: 0
 Query Match: 58.89% Indels: 2
 DB: 9 Gaps: 2

US-09-970-944-2 (1-899) x BC009333 (1-2688)

QY 359 AspValAlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeu 378
 Db 2 GAGTGGCCCTCATGTGGGCGCTCATCGCGTGGCGCTGCTGCTGCTGCTGCT 61
 QY 379 ValLeuIleLeuValTyrCysArgLysLeuGlyLeuGlyLeuAspSerAspValAlaAspSer 398
 Db 62 GTCTCATCTCTGTTATTCCTCCGGAAGAGAGGGGCTGGACTCAGATGTGGTGAATCG 121
 QY 399 SerIleLeuThrSerGlyPheGlnProValSerIleLeuLysProSerLysAlaAspAsnPro 418
 Db 122 TCCATTCTACCTCAGCTTCAGCGCTCAGCATCAAGCCGACCAAGCAGCAACACCCC 181
 QY 419 HisLeuLeuThrIleGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThr 437

Db 182 CATCTGCTCACCATCCAGCGGAGACCTCAGCACCAACCACCACTTACCAGGGCAGTCTC 241
 QY 438 CysProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSer 457
 Db 242 TGTCCCGGCGAGATGGGCGCCAGCCCAAGTTCAGCTCACCATAATGGGACCTGTCTCAGC 301
 QY 458 ProLeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGlu 477
 Db 302 CCCCTGGGTGGCGCGCGCCACACACTGCACCACTCTCCCACTCTCTGAGCGCGAGGAG 361
 QY 478 PheValSerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsn 497
 Db 362 TTCGTCTCCGCGCTCTCCACCGAAGACTTCTCCCTCCCTCCCGGAGGACCAAGCAAC 421
 QY 498 MetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIle 517
 Db 422 ATGACCTATGGGACCTTCAACTCTCTCGGGGCGCGGTGATGATCCCTAATACAGGAATC 481
 QY 518 SerLeuLeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThr 537
 Db 482 AGCCTCTCTATCCCCCAGATGCCATACCCGAGGAGAGATCTATGAGATCTACCTCAGC 541
 QY 538 LeuHisLysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerPro 557
 Db 542 CTGCACAGCCCGAAGACGTGAGGTTCCTCTAGCTGGCTGTACAGACCTCTGTAGTCTCC 601
 QY 558 IleValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAsp 577
 Db 602 ATCGTGTAGCTGTGGACCCCTGGCGTCTCTGTCTACCCCGCCAGTCATCTCGCTATGAC 661
 QY 578 HisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGlu 597
 Db 662 CACTGTGGGAGCCCGACCCCTGACAGCTGGAGCCTGGCCCTCAAAAAGACGTCGTGCGAG 721
 QY 598 GlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyr 617
 Db 722 GGCAGCTGGAG---GATGTCTGACCTGGCGGAGGAGGCGCCCTCCACCTCTACTAC 778
 QY 618 CysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeu 637
 Db 779 TGCCAGCTGGAGGCGCAGTGTCTGCTCTTCCACCGAGCAGCTGGGCGCTTTGCCCTG 838
 QY 638 ValGlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuGlnGluPro 657
 Db 839 GTGGAGAGCCCTCAGCGTGGCTGCGCGCAAGGCGCTCAAGCTGCTTCTGTTTGGCGCG 898
 QY 658 ValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAsp 677
 Db 899 GTGGCTGCACTCCCTCGAGTCAACATCCGGGTCTACTGCTGCTGATGACACCCAGAT 958
 QY 678 AlaLeuLysGluValValGlnLeuGlyLysGlnLeuGlyGlyGlnLeuIleGlnGluPro 697
 Db 959 GCACCTAAGAGAGTGGTGCAGCTGGAGAACAGCAGCTGGGGGACAGCTGATCCAGAGCCA 1018
 QY 698 ArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValPro 717
 Db 1019 CGGTCTCTGCACTTCAAGGACAGTTACCAACACTGCGCTTATCCATCCACGATGTGCC 1078
 QY 718 SerSerLeuTrpLysSerLysLeuValSerTyrGlnGluIleProPheTyrHisIle 737
 Db 1079 AGTCTCCTGTGGAGAGTAAAGCTCTCTGTGAGTACAGAGATCCCTTTTATCATC 1138
 QY 738 TrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSer 757
 Db 1139 TGGAAATGGCAGCAGCGGTACTTGCATCCACTTCCCTTGAGCGGTGTACGCCCCAGC 1198
 QY 758 ThrSerAspLeuAlaCysLysLeuTrpValGlnValGluGlyAspGlyGlnSerPhe 777
 Db 1199 ACTAGTAGCTGGCCCTGCAAGCTGTGGGTGTGGAGGTGGAGGCGACGGGAGAGCTTC 1258
 QY 778 SerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuAlaLeuGluSer 797
 Db 1259 AGCATCACTTCAACATCAACAGGACACAAGGTTTGTGCTGTGCTGTGCTGTGAGAGT 1318

QY	798	GluAlaGlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArg	817	GRGYQKRTTRCTNPAPLNGGAFCEGQSVQKIACTTTLCPVDGRWTSKWSKSTCGTECH
Db	1319	GAAGCGGGGTCCAGACCCCTGGTGGGCCAGTCCCTTCAAGATCCCTCTCATTCGG	1378	WRRECTAPAKNGKDCDGLVQSKNCTDGLCWAAPSDSDVALYVGIIVAVTVCLIA
QY	818	GlnLysIleIleSerSerLeuAspProProCysArgGlyAlaAspTrpAtgThrLeu	837	ITVVAVLVYRKNHRDPSDIIDSSALNGCFQPNIKARQDILLAVPDDLTLSAAVMYR
Db	1379	CAGAAGATAATTTCCAGCCTGGAGCCACCTGTAGGCGGGTGCAGACTGGCGACACTG	1438	GPVALHNSDKIPMTNSPILDEPNLIKIVYSSGAVTPQDDLLAEPSSKLSPMOTKS
QY	838	AlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThr	857	LENEALNKNQSLAQRTDPSCTAFGTFFSGGLHLIIPNSGVLSLLIPAGALPQGRVYE
Db	1439	GCCCAGAAATCCACCTGGACGACATCTCAGCTTCITTGCTCTCAAGCCAGCCACACA	1498	MYTVHRKNMRPMDSDQTLTPVWSCGPGGALLTRPVLTLHHCADPSTEDWKIQ
QY	858	AlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeu	877	KNQAVQOMEDVVVGBENFTTTCYIQDAAECHILTENLSTVALVQSTTKAAAKEL
Db	1499	GCCATGATCTCAACCTGTGGAGCGCGGCACTTCCCAACGGCAACCTCAGCGACCTG	1558	KLAIFGLCCSSLEYSIRVYCLDDTDALKEVLQEROMGQQLLEKPEKALHFVKISIN
QY	878	AlaAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAla	897	LRLSIHDIASLWKSLLAKYQRIPEFHIWSGSORNHLCTFTLERLSLNTVLCKLC
Db	1559	GCTCAGCAGTGGTGGACTGGCGCAGCCAGCAGCTGGCCCTCTTCACAGTGTGGAGGCT	1618	VROVEGGQIFQLNCTVSEPTGIDILPLDPASTITVTGFSAPFSIPLPIROKLCSLI
QY	898	GluCys 899		DAPQTRGHDWRMLAHLKLNLDRLYNYFATKSPSTGVLIDLWEAQNFPGNLSMLAAVLE
Db	1619	GAGTGC 1624		EMGRHETVVSAAEQY"
RESULT 12	MMU72634	9299 bp mRNA linear ROD 13-MAY-1997		
LOCUS	MMU72634	Mus musculus rostral cerebellar malformation protein (rcm) mRNA,		
DEFINITION	U72634	complete cds.		
ACCESSION	U72634.1	GI:2088526		
VERSION	U72634.1	GI:2088526		
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
gene				
CDS				


```

LOCUS      AB118026      9328 bp      mRNA      linear      ROD 19-DEC-2003
DEFINITION Rattus norvegicus Unc5h3 mRNA, complete cds.
ACCESSION  AB118026
VERSION    AB118026.1 GI:40217509
KEYWORDS   .
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1
AUTHORS    Kuramoto, T., Kuwamura, M. and Serikawa, T.
TITLE      Rat neurological mutations cerebellar vermis defect and hobble are
            caused by mutations in the netrin-1 receptor gene Unc5h3
JOURNAL    Mol. Brain Res. (2003) In press
REFERENCE  2 (bases 1 to 9328)
AUTHORS    Kuramoto, T. and Serikawa, T.
TITLE      Direct Submission
JOURNAL    Submitted (22-AUG-2003) Takaishi Kuramoto, Institute of Laboratory
            Animals, Graduate School of Medicine, Kyoto University;
            Yoshidakonoe-cho, Sakyo-ku, Kyoto 606-8501, Japan
            (E-mail: tkuramoto@anim.med.kyoto-u.ac.jp,
            URL: www.anim.med.kyoto-u.ac.jp, Tel: 81-75-753-4494,
            Fax: 81-75-753-4409)
FEATURES   Location/Qualifiers
            1..9328
                /organism="Rattus norvegicus"
                /mol_type="mRNA"
                /db_xref="taxon:10116"
                /chromosome="2"
            gene     1..9328
                /gene="Unc5h3"
                /name="Unc5h3"
                /codon_start=1
                /protein_id="BA005181.1"
                /db_xref="GI:40217510"
            CDS      151..2946
                /translation="MRKGRATAPAAAGLGLGLYLQMLVLPALALLSAGSGTSAOQDD
                FFHELPTFSPDPPELPHLEPIERAYIVKNKPNVLYCKASPATOIYFKCNSEVHQ
                KHWDERVDETSGLVIREVSIERSQVEELGPDYWCQCVAMSGAGTKSRKAYV
                TIHNLILIQARLSDTIVAKNIVAKRKSTTATVIVYNGGNTWAEWSCNRC
                GRGQRKTRTNPALNGAFCEGVQKIACTILCPDVRWTSKWSKSTCTGCTH
                WREKCTAPKNGKDCGLVLQSKNCTDGLCWAAPSDDDVALVGVIAVTGCLA
                ITVVVALFVYRKNRHPFESNIIDSALNGGFQVNIKAARQDLAYPPDLTSAAMYR
                GPVVALHDVSDKIPMTNSPTLDPLNKKIKNVSSGNVTPQDLAFSSKLSFQMTQS
                LLENEALNKNQSLAQTPSCFAGTFNSLGGHLIPNSGVSLIIPAGAI PQGRVYE
                MYTVIRKENRPFMDSDQLLPVVSQGPFGALLTRPVILTHHCADPNTEDWKIQL
                KNVAVQGWEDVVVVEENFTPCYIQLDAEACHILTENLSTVALVQSTKAAAKRL
                KLAIFGLCCSLEYIRVYCLDQDALKVELQEROMGQLLEPKALHFKGSITH
                LRLSIHDI THSLWKSLKAKYQEI PFVHIWSSGORNLCFTFLERISLNTVELVCKL
                VRQVEGGOIFOLNCTVSEPTGIDLPILDPASTITVTGPSAFSIPILPIRQKLCSSL
                DAPOTRGHDWRMLAHKLNLDRLYNLPATKSSPTGVILDLWEAQFPDGNLSMLAAVLE
                ENGRHETVSLAAGGY"
            2947..9328
                /gene="Unc5h3"
            repeat_region 8727..8772
                /note="simple sequence repeats"
            polyA_signal 9309..9314
                /rpt_unit="gt"
                /gene="Unc5h3"
            ORIGIN
Alignment Scores:
Pred. No.:      7,74e-161      Length:      9328
Score:          2688.50        Matches:     512
Percent Similarity: 72.95%      Conservative: 154
Best Local Similarity: 56.08%      Mismatches: 216
Query Match:    56.16%         Indels:      31
DB:             10             Gaps:        12

US-09-970-944-2 (1-899) x AB118026 (1-9328)

```

```

QY      9  ProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSer 28
      226  CCTGCCCTG-----GCCCTGCTCAGGCCAGTAGTCACCGCTCGGCCGCTCAAGATGAT 279
QY      29  -----AlaThrValAlaAsnProValProGlyAlaAsnProGlyAlaAsnProAspLeuLeuProHis 45
      280  GACTTTTTCAGTAACTCCCAAGAAACCTTCCATCAGACCACCTGAGCCATTCACACAC 339
QY      46  PheLeuValGluProGluAspValTyrIleValIleValIleValIleValIleValIleVal 65
      340  TTTCTCATTTGAGCCCGAGGAAGCTTACATTTGGAAGAACAAAGCCTGTGAACCTGTATTGT 399
QY      66  LysAlaValProAlaThrGlnIlePhePhePhePhePhePhePhePhePhePhePhePhePhe 85
      400  AAGCCAGCCTGTCACCCAGATCTACTTCAAGTGAACACGCAATGGGTTTCATCAGAAG 459
QY      86  AspHisValIleGluArgSerThrAspGlySerThrAspGlySerGlyGluProThrMetGluValArg 105
      460  GACCACGTAGTAGACGAGAGAGGAGTAACTCCCGTCTCATTTGTGAGAGAAAGTGAAGC 519
QY      106  IleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrpCysGln 125
      520  ATTGAGATTTCAGCCAGCAGGTGGAGAACTGTTCGGCCTGAAGATTACTGTGTGCAG 579
QY      126  CysValAlaTrpSerSerSerGlyThrThrIleSerGlnIleValIleValIleValIleVal 145
      580  TGTGTGCTTGGAGCTCAGCGGCGACTACGAAAGTCCGAAAGTCCGAAAGTCCGCGCATTCGC 639
QY      146  ArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGluGlnGly 165
      640  TATCTCGGGAAGACATTTGAGCAGGAACCTTGGGAAAGAGTGTCTTGGAGCAGAA 699
QY      166  IleValLeuProCysArgProGluGlyIleProProAlaGluValGluTrpLeuArg 185
      700  GTGTTACTCCAGTGGCGGCCACCTGAAGGATCCCAATGCTCGAGGTGGATGGCTTAAG 759
QY      186  AsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSer 205
      760  AATGAAGACATAATTGATCTCTGGAAGATCGGAATTTTTATATCATATCATCATCAAC 819
QY      206  LeuValAlaArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsn 225
      820  CTGATCATCAAGCAGGACGACTCTCAGATACAGCAATATATACCTGTGTGTGCCAAAAAC 879
QY      226  IleValAlaArgArgSerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTyr 245
      880  ATTTGTTGCCAAGAGGAAAAAGCACACAGCCACCGTCATCGTGTATGTATTATGTGCTCG 939
QY      246  SerThrTrpThrClnTrpSerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArg 265
      940  TCACCTGGGCGAATGGTCTGTGTGTAAACAGCCGCTGTGGCGAGGATATCAGAAACGC 999
QY      266  SerArgSerCysThrAsnProAlaProLeuAsnGlyValAlaPheCysGluGlyGlnAsn 285
      1000  ACAAGAACCTGCACCAACCCGCCCTCATTTGTTGGGGCTTCTGCAGGGGCGAGAT 1059
QY      286  ValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrpSerProTyr 305
      1060  GTA---CAGAAAATAGCATGCACCGTTATGCCAGTGGATGGAGGTTCGTGTGG 1116
QY      306  SerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCysSerAsp 325
      1117  AGCAATGGTCAACTTGTGGGACTGAGTGCACCCACTGGCGAGGAGGAGGTATCAGCA 1176
QY      326  ProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCys 345
      1177  CCAGCCCCCAAGAACGGGGTAGAGCTGTGATGCGCTGCTCTCAATCCAAAGAACTGC 1236
QY      346  ThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyrValGly 365
      1237  ACCGATGGGCTGTGCATGCAGGCTGCTCTCTGACTCAGATGACGTGCTCTCTACGTTGG 1296
QY      366  Leu---IleAlaValAlaValCysLeuValLeuLeuLeuLeuLeuLeuValTyr 384

```


Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

JOURNAL

REFERENCE

Submitted (01-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 200968)
 Biren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kellis,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mieng,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ranasingh,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupbach,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

JOURNAL

COMMENT

Direct Submission
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L26431
 Center clone name: 335_A.11
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 136489 bases at least Q40
 Consensus quality: 199303 bases at least Q30
 Consensus quality: 199583 bases at least Q20
 Insert size: 197000; agarose-fp
 Insert size: 200168; sum-of-coverage
 Quality coverage: 9.9 in Q20 bases; agarose-fp
 Quality coverage: 9.7 in Q20 bases; sum-of-coverage

 * NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
 * 1 44442: contig of 44442 bp in length
 * 44443 44542: gap of 100 bp
 * 44543 45239: contig of 697 bp in length
 * 45239 45339: gap of 100 bp
 * 45340 46467: contig of 1128 bp in length
 * 46468 46567: gap of 100 bp
 * 46568 55518: contig of 8951 bp in length
 * 55519 55618: gap of 100 bp
 * 55619 122673: contig of 67055 bp in length
 * 122673 122773: gap of 100 bp

* 122774 137362: contig of 14589 bp in length
 * 137363 137462: gap of 100 bp
 * 137463 162118: contig of 24656 bp in length
 * 162119 162218: gap of 100 bp
 * 162219 187047: contig of 24829 bp in length
 * 187048 187147: gap of 100 bp
 * 187148 200968: contig of 13821 bp in length.

FEATURES

source
 1. .200968
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP23-335A11"
 /clone_lib="RPC1-23 Female Mouse BAC"
 1. .44442
 /note="assembly_fragment"
 clone end:SP6
 vector_side:left
 44543..45239
 /note="assembly_fragment"
 45340..46467
 /note="assembly_fragment"
 46568..55518
 /note="assembly_fragment"
 55619..122673
 /note="assembly_fragment"
 122774..137362
 /note="assembly_fragment"
 137463..162118
 /note="assembly_fragment"
 162219..187047
 /note="assembly_fragment"
 187148..200968
 /note="assembly_fragment"
 clone end:T7
 vector_side:right

ORIGIN

Alignment Scores:
 Pred. No.: 1.27e-158 Length: 200968
 Score: 2681.00 Matches: 771
 Percent Similarity: 25.66% Conservativeness: 28
 Best Local Similarity: 24.76% Mismatches: 42
 Query Match: 5 Indels: 2278
 DB: 2 Gaps: 18
 US-09-970-944-2 (1-899) x AC123700 (1-200968)
 QY 59 LysproValLeuValCysLysAlaValProAlaThrGlnIlePhePhe----- 75
 Db 101981 AGCCATACCTGCTTACATCCAGAGATAAACACGACGATCCCTCGAGTTCCTCCAGGC 101922
 QY 76 -----LysCysAsnGly-----GluTrpValArgGlnValAspHisVal 88
 Db 101921 CTCGGAGGCTACAGGGTCTATAAGCTCTAGTGTGGGGGGGCGATCAACCATTTG 101862
 QY 89 IleGluArgSerThrAsp-----GlySerSerGlyGluProThrMetGluValArg 105
 Db 101861 GGACGGGAGACTCAGTGTGGCTCTGAACCCAGCAGAGATTGCCAACCATGGAGGTCCGG 101802
 QY 106 IleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGlyTrpCysGln 125
 Db 101801 ATCAACGTATCAAGGCAGGAGTTCGAGAAAGTGTGGCTCGAGAGTACTGGTCCAG 101742
 QY 126 CysValAlaTrpSerSerGlyThrThrLysSerGlnLysAlaTrpIleArgIleAla 145
 Db 101741 TGTGTGGCATGGAGGCTCTCTCAGAAACACCAAAGCCAGAGGCTTACATCCGGATTGCC 101682
 QY 145 ----- 145
 Db 101681 TGTGAGTACTGGACCGAGGCTGTGGCATCCGGGCGATGCTACTAGAGGGTATCCCTGGA 101622
 QY 145 ----- 145

Db	101621	GAGAGGGCGGCAAGCTTCCTGGAGGGGTAGAGGCAGTGCCTTAGGAACATAAATCACCCAC	101562	Db	100543	GAGCAGCCCACTCACCCGAGCTTATAAGAACTTTGCCCTCTCTGCCCAATCCCTTGGCCC	100484
QY	145	-----	145	QY	242	-----	Asn 242
Db	101561	ACTCCTCAATGCACATGGTGCCTCCTTCGGCCATGACGTGTGCATGACGTCTGCAGGC	101502	Db	100483	AGCATGCTGAGGAAGCCCGCTCCTGACCCAGTCTCTCTCTGTCGGCCAGTGAAC	100424
QY	146	-----	ArgLeuArgLysAsnPheGlu	QY	243	GlyGlyTyrSerThrTyrThrGluTyrSerValCysSerAlaSerCysGlyArgGlyTyr	262
Db	101501	CAAGAGTCAGCTGATCCAGACCCCGGTTTCTCCTCCAGATTTCGCAAGAACTTTGAG	101442	Db	100423	GGTGGGTGTCGACGTGGACCGAGTGGTGGTCTGTCAGTCCAGCTGTCGGCGTGGCTGG	100364
QY	153	GlnGluProLeuAlaLysGluValSerLeuGluGlnGlyValLeuProCysArgPro	172	QY	263	GlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGlu	282
Db	101441	CAGAGCGCTGGCCAAAGGAAGTGTCACTGAGCAAGCAATGTGCTTACCTTGTGCGCCC	101382	Db	100363	CAGAAACGAGCGAGCTGCACCAACCCGACCTCTCAACGGGGCGCTTCTGTGAG	100304
QY	173	ProGluGlyIleProProAlaGluVal	181	QY	283	GlyGluAsnVal	286
Db	101381	CCGGAGGANTCCCCCAGCTGAGGT-GAGTGGAGGGTGGGGCTTCCAGACCTCTATC	101323	Db	100303	GGCGAGAATGTCAGAAAACAGCTGGGCCACCTCTGTGCCAGGTACCGGAGCCACCGA	100244
QY	181	-----	181	QY	286	-----	286
Db	101322	CTACCTAAAGCTCAGACTCCTGGTGCACCTCCCGAGTGTGCTCTATACCTGTTC	101263	Db	100243	CTCTGTCATCGTTCTCAACACGCCATCTCCGTGGCTGGCTCCATCGCACCCACACAGCC	100184
QY	182	-----	GluTyrLeuArg	QY	286	-----	286
Db	101262	GGCCAGCACGGGGCCAGCGGTGTGTGTACCGCTTCTCCCGCTAGTGGAGTGGCTCGA	101203	Db	100183	TGCCCGCCCTCCTCCCATGGCTCCATCCACCCGCTGCCACACACAGGCGCCAGGTCA	100124
QY	186	AsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSer	205	QY	286	-----	286
Db	101202	AATGAGGACCTCTGTGGACCTCTCCTCGACCCCAATGTGTACATCACAGGGGAGCAGC	101143	Db	100123	GTCGTAGGTGGGATCAGGTGTCAGTGTGTGTCAGATCAAGGGCTCGGTTAGGGTCTAA	100064
QY	206	LeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsn	225	QY	286	-----	286
Db	101142	CTAGTCGTGCGCGAGCGCGCTGCGGCACACTGCCAACTACCTCGCTGGCGCAAGAAC	101083	Db	100063	GGTCAGTTTTAGSACAGGATCTAGGCTCGCGCTGGGGCTGCACAATCCACCTGTGCCTG	100004
QY	226	IleValAlaArgArgArgSerAlaSerAlaValIleValTyrVal	241	QY	286	-----	286
Db	101082	ATCGTGGCCCGTGGCGAAGCGCTCTCGCGCCGTCATTGTTATGG-TGGGTCTGGGC	101024	Db	100003	GACCAGAAGGCCCATCATGTGACTAGGAGCCCTTACCGCCTCAGACAGGAGACAGCTCC	99944
QY	241	-----	241	QY	286	-----	286
Db	101023	TCCCAGTCACAGGAAGAGGTACTGAGGTGCCAACAGGGAATGGTGTGTGGCTCAGAGG	100964	Db	99943	CTGGTTGAGAGCGCCGAGGAGCCCGCCAGCATCTTGGCTCTCTTGTCTATTAGCATGTG	99884
QY	241	-----	241	QY	287	-----	HisAspArgThrValSerSer
Db	100963	ACGGAAGGCATGGAGCAAGAACACAGGCTAACAGACAGTGGCTTGAATAAGAAAGAG	100904	Db	99883	TCATCAAGGCTAGCTGGACACTCAAGTCCGACCGATCAGAGGGCTCACCTTGGCATTTG	99824
QY	241	-----	241	QY	293	-----	293
Db	100903	CAGAAAGTGTATCTGGGCAAGGATGGGACATGTGGTCACTTGTGAGTGGGAGCATGG	100844	Db	99823	TACCAAGCCCGAGCCCTGCTGTGTCCAGATGACTAGCGGGGCCACCCAGCGCCAGCCACC	99764
QY	241	-----	241	QY	294	-----	LeuLeuValSerVal
Db	100843	TGCAAGAGAGCAAGAGACAGAGGAGGCTCAGAGGAGCGTAACATCAACCTTGGCGGT	100784	Db	99763	CAGGCAACCTCTTAAGCTGACTGTCCATTTCCCTCTCTGCTGTCTCTTGTTCAGTG	99704
QY	241	-----	241	QY	299	AspGlySerTyrSerProTyrSerTyrSerAlaCysGlyLeuAspCysThrHisTyr	318
Db	100783	CTCAGTCCCCGAAATGGGAGAGAGCGCTCAGAAATGAGACGCTTCTCTCCCCCAAGAGCA	100724	Db	99703	GATGGAGCTGGAGCCCATGAGTAAAGTGGTCAAGCTCGGGCTTGAAGTCCACCCACTGG	99644
QY	241	-----	241	QY	319	ArgSerArgGluCysSerAspProAlaProArgGlnGlyGluCysGlnGlyThr	338
Db	100723	GAAGAGGTGGATAGGAGCTGCCCTGCGAAGAGTGGGTTCTAGCTGGAGGTGCCCTGT	100664	Db	99643	CGGAGCCGGAGTGTCTCGACCCCGCCCGAACCGAGGTGAGAGTGCCTGGGGGTGCT	99584
QY	241	-----	241	QY	339	AspLeuAspThrArgAsnCysThrSerAspLeuValHis	352
Db	100663	GACCTGTAGTTTGTCTGGCCACCTTGACACTTGGCCAGAGTAGATATCTGAAGGCTCT	100604	Db	99583	GACCTGGACACCCGCAACTGTACAGTACCTCTGCTGCACAGTGTCTCTCTGCCCC	99524
QY	241	-----	241	QY	352	-----	352
Db	100603	GCCAGCCCTGGGGCCAGCAGCCTAGGTTGGTCTCTCAGGGTTTCTCCAGGCACAG	100544	Db	99523	CAGGGCTCACCTCCCTCTGACAGCCCTATGGGCTGTGGTGTGAGGAGCCACTCTCTGCCC	99464
QY	241	-----	241	QY	352	-----	352
Db	100544	-----	241	Db	99463	TGTGTCTCTCCAAAGCTTGGGAGCCAGGGGTCCCTTTTCATCTCTAGCACACTCTTCCCT	99404

QY 352 ----- 352
Db 99403 GTTACACAGACACACCCTTCCCTCGAGACATGCGATGTGAATCTTGACCCCTGT 99344
QY 352 ----- 352
Db 99343 ACATCCAACTGCTCCCGCGGCTTTGAAATTATCCCGTTCCCAAGCTCACACATAAGT 99284
QY 352 ----- 352
Db 99283 CCTGTGGGGGTGCTGGGCTTGTGAGCGCTGTTAATTACAAGGCCATAACTAACAAATAAA 99224
QY 352 ----- 352
Db 99223 GATCTCTGTTGGAATGTTTTTCTTTAGACTACACCTCCCGTGTCTTTCAATTAAA 99164
QY 352 ----- 352
Db 99163 AACCTAATTATCTGTGTCATTCCTTTCTGATGACATTTCTTCCATTTATGCAAAACTCT 99104
QY 352 ----- 352
Db 99103 TAAAAATGATAGTCTTCTCCAAGCAGTTGGTCCCGCCACCCACCCAGTTTCAGGGCT 99044
QY 352 ----- 352
Db 99043 GGTGGTATACAAAGGCTCACACTGTCTGGAACACCTGAGAACTGACCACCCATGT 98984
QY 352 ----- 352
Db 98983 ATCTGCCATTTTCTATTCTATAATGAACCTAGCAAAAGAAATTTATTAGTTTTAGAGGC 98924
QY 352 ----- 352
Db 98923 TCAAATTCAAATAGCTGATGTGGCTCTGGTAAATGTGCTTTTACCCCAAGCCACAT 98864
QY 352 ----- 352
Db 98863 TTCTGTGGCAGATAGCACCATGTCTAGTACTGTAGACAGAGAGATCACGTGTGAGG 98804
QY 352 ----- 352
Db 98803 AAGAGAAAAAGACAGACAGACAGAGAGGTAGACACGGAAGATGGACAGCCCATGTCT 98744
QY 352 ----- 352
Db 98743 AAGAGCAGCACTCATATGATGTAAATGCTCTCCCACTGGGGCCACCTCCTAAAAGGCC 98684
QY 352 ----- 352
Db 98683 CACTGTCTCATATTGCCAATGGGGACCAACACCTGTCAATATGATCCTTTAGGCAGG 98624
QY 352 ----- 352
Db 98623 CATGTGCAAACTATAGCACCATACTTAGGGTCTCTCAGGAGTTTACCTTGCCCGCAGCAC 98564
QY 352 ----- 352
Db 98563 CCTGACCCCGAGCCTAGGAGCCAGGATATACCACTCACCCCGGTGAAGTACTCT 98504
QY 352 ----- 352
Db 98503 CCCCTGTCTCACTCTGCTGTGAGGTGTGACTTGTCTCAATCCCAAGCAAGGTGTGGTG 98444
QY 352 ----- 352
Db 98443 GGTATAGAACTAGGATGTACCCACCAATACCAAGCTCCCTGCCACGTGCCCACTGCCAG 98384
QY 352 ----- 352
Db 98383 GGGGTGCTGAGATCCTATTCTCTGGGGATAAACCATGGGCTCTCTCTGAAGCCTCTA 98324

QY 353 ----- Ser-AlaSerGlyProGluAspVa 360
Db 98323 CATCCCTGCCANTCTCAAGACACACTTTCTTTTCAGCCCTCTTCCGGCCCCGAGGAGCT 98264
QY 360 lAlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValIle 380
Db 98263 GGCTCTCTACATCGGCTCTGTCGGCGTGGCTGCTCATCTCTGCTGCTGCTGCTGCT 98204
QY 380 uIleLeuValTyrCysArgLysLysGluGlyLeuAspSerValAlaAspSerSerIle 400
Db 98203 CGTCTCATCTACTCTCCGCAAGAGGAGTGGACTCAGACTGAGCTGGCTGACTCATCCAT 98144
QY 400 eLeuThrSerGlyPheGlnProValSerIleLysProSerLysAla 415
Db 98143 CCTTACCTCAGGCTTTCAGCCCTGTGAGCATCAAGCCAGCAAGCAGGTGAGGTGCCACC 98084
QY 415 ----- 415
Db 98083 TACCCCTGCTCTCCTGCTGAGCACCCCAAGCAAGGCTGTTGATGGGGAGTGCCTA 98024
QY 416 ----- AspAsnProHisLeuLeuThrIleGlnProAspLeuS 428
Db 98023 CCTGAATCCATCTTTATCTCTGTAGACAATCCCATCTGCTCACCATCCACCGACCTCA 97964
QY 428 er---ThrThrThrThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProL 447
Db 97963 GCACCACGACACCACTACAGGGCAGCCTGTCTCCCGCAGAGTGGACCCAGCCCA 97904
QY 447 YsPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyArgHisThrLeuH 467
Db 97903 AGTTCAGCTCTCTAATGGTCACTGCTGAGCCACTGGGAGTGGCGCCCATACGCTGC 97844
QY 467 iHisSerSerProThrSerGluAlaGluPheValSerArgLeuSerThrGlnAsnT 487
Db 97843 ACCACAGTCCCCACCTCTGAGGCTGAGGACTTCGTCTCCGCTCTCCACCCAAAAT 97784
QY 487 YrPheArgSerLeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuG 507
Db 97783 ACTTTCCTTCTCTGCCCCGGGTACCAGCAACATGGCCTATGGGACCTTCAACTTCTCG 97724
QY 507 lyGlyArgLeuMetIleProAsn----- 514
Db 97723 GGGGCCGCTGATGATGCCCTTAAACAGGTAGGCCCCCTAGAACTCCCGAGAGAACATGAGA 97664
QY 514 ----- 514
Db 97663 GGTAAAGACTAGTCTTGTGTCATCTTTTCTGCCAGAGTGACACGCTCGTGCAGAGTGGCC 97604
QY 514 ----- 514
Db 97603 CAGTCTGGCCAGTATAGGAAGGATGAGTTGGCGCTATAGCAAGCTGTGAAAGGGTC 97544
QY 514 ----- 514
Db 97543 TCCTTGTGTGTTGGACACGAAAGGATTGATGTTTCAATTATCACACAGTACAGCTTAC 97484
QY 514 ----- 514
Db 97483 ATGTTGATAGGTGCTGTCATCGGCGCCCTAACACAGACCCCGAGAGGTTCACAAACAAGAC 97424
QY 514 ----- 514
Db 97423 AGCAGAGCACCTTTAGAACCTGCGCCCTCCAGCGTCTCTCCATCTCATCCACCACTCCA 97364
QY 514 ----- 514
Db 97363 CAGAAAGTGCCTCAGACAGGGTCCCTGAGACCTTCTTGTAGAGATCCACAGACCCCTCCA 97304
QY 514 ----- 514
Db 97303 AGGCCATCTTTGGAGAAATCTCTGCAACCTCTGACAAGGCTGACCAAGGCTGACCCCACTCTGACAT 97244
QY 514 ----- 514

Db	97243	TCTCTGTTCCTCTCCTT	AGGTGACCCCA	CATATCTCCTGTTC	TGTGTCCTGACCATCCCTGGCTGT	97184
QY	514	-----	-----	-----	-----	514
Db	97183	TCCGGTATAGCGGTTCT	CCTCAGGATACCCCT	TCCCAACCCACCCCACTTC	CCGCCGTGC	97124
QY	514	-----	-----	-----	-----	514
Db	97123	CAITTTCTGAGTATCT	CACCTACTCTCTCCAT	CCAGCCGCCAGTTCT	TTGCAGAGCATGG	97064
QY	514	-----	-----	-----	-----	514
Db	97063	AACTCTGCATTCAGCT	CGCTCTTTGGACAGTGT	CAITTCAGCTCCANCT	GCCTGTCTCAGACACCTCGGAG	97004
QY	514	-----	-----	-----	-----	514
Db	97003	TCITTTGTGCCGMAA	CAGAACCGTCTCCAC	CCCTTTCTCTCAGCTAC	CTCTCCTCGGCT	96944
QY	514	-----	-----	-----	-----	514
Db	96943	TTCCCCGGCTCACGG	TTGGCAGAGTGTCAIT	TGCTTAAGCTAGAACT	GGCGTGCAC	96884
QY	514	-----	-----	-----	-----	514
Db	96883	CTTGGGCTCCTGTCT	TTTCTATTCCTATTC	CCCCCTCGCTATTA	AGTCAGCTGTATCATGAT	96824
QY	514	-----	-----	-----	-----	514
Db	96823	GTTGCATGCTCCCT	TTCCCAATAGCTCCT	GTCTCCTCTGCTCT	CCTCATCCTGCCCTGGT	96764
QY	514	-----	-----	-----	-----	514
Db	96763	GCGTCACACACAC	CGGCTAGCAGCTCCT	TTACCTCAACCGCG	CACTACACATTTACT	96704
QY	514	-----	-----	-----	-----	514
Db	96703	GAATGAATCTAGAG	TTGATTAAGCCCCCT	CCCCACTCAATCCCT	TCTAAGTTCCAGTT	96644
QY	514	-----	-----	-----	-----	514
Db	96643	GCTCTCTACAGAG	CCCCAGTGACAAGG	TAATGCAACCCCACT	GTGCTAGCCGCT	96584
QY	514	-----	-----	-----	-----	514
Db	96583	GTGTGCTTTTCCCT	CCTATATGCACTT	CTGTATGTGACAGCC	CTTACAGACCTTTGGCAC	96524
QY	514	-----	-----	-----	-----	514
Db	96523	CTGAACATGACACT	GTCCCTGGCTCAT	AAGTAATGCTCTT	TCTACCTTTCTCCTGGC	96464
QY	514	-----	-----	-----	-----	514
Db	96463	TCATCTGTTGTGCT	GTGCCTCATACACC	CTCAGGACAGGGCAT	GAGAAATCTC	96404
QY	514	-----	-----	-----	-----	514
Db	96403	ATAGCTAGTGAAGT	AGCTGGTCCCTGTT	AGCTGTATACAAAA	AGGCAGACAGATGGTG	96344
QY	514	-----	-----	-----	-----	514
Db	96343	GTAATCCGTGGCC	CTGCAGTGTGCC	ACCCTGTGTCTCT	CTGTGTCTAGTTGCCT	96284
QY	514	-----	-----	-----	-----	514
Db	96283	CTCTCCTTGTGACT	GAGAAAGCTAGG	CCCTTTTCTCTTAG	CTTTCTGAGTGGCA	96224
QY	514	-----	-----	-----	-----	514
Db	96223	GCATGAGTATTCA	ATCTTTGAATPACT	CTCTCCAGCAGG	CAITTCACACGAGGTTGTG	96164
QY	514	-----	-----	-----	-----	514

Db	96163	ACTGCCCCGACTCC	AGGAGCATCAGGTT	CTTACGGTTCAGCACAG	GTTAGTGAAGAAGC	96104
QY	514	-----	-----	-----	-----	514
Db	96103	AAAAGATTTCAG	CACATGGGTTTC	GAGCTGTCTGTGAC	AGCTCGGTGTAAGACCCATC	96044
QY	514	-----	-----	-----	-----	514
Db	96043	ACTCGGTGTAAG	ACACCGGACAGGT	ATACATCTGTAGAC	ACACTTAACCACTCGACGC	95984
QY	514	-----	-----	-----	-----	514
Db	95983	TTCACTGTCCG	GAGGCTGTGTGAC	CCCTATCGAACCAT	CACTACTGTGAGTTTGTAGAATATAA	95924
QY	514	-----	-----	-----	-----	514
Db	95923	GTGTCTTACAT	AGCCCTTGAACAG	CGCTGTGAGTTC	CCAGCCAGCAACACAGTGGAGCTATC	95864
QY	514	-----	-----	-----	-----	514
Db	95863	TGCTGTTTTAG	GGCAAGTTCAG	ATGTCTCCCCCT	TTCTGCCCTCAGAGCAGGACC	95804
QY	514	-----	-----	-----	-----	514
Db	95803	TCAATGCTGCT	CTCCGTTGATCT	CAGCCCTTGTCT	CTCCACTACAAAACCTCCC	95744
QY	514	-----	-----	-----	-----	514
Db	95743	CTTCTCCACAC	ACTGATACCACT	CTCTGGGCAAGCG	CGATCAACAGACCGAGGAGACA	95684
QY	514	-----	-----	-----	-----	514
Db	95683	TTGGAGCGGTAG	ATCTTTGGCCCT	TGAGGCTATTAT	AGACCTGCCCCCTCCCGCAAG	95624
QY	514	-----	-----	-----	-----	514
Db	95623	GATGGACAAAT	GACAGGAGAGAC	CCAGTTCTT	TAGAGGGGACAGGACTTAAGGGGGGAGGA	95564
QY	514	-----	-----	-----	-----	514
Db	95563	GCTTTAGCAAG	ACAAAGGCCCT	GTCTGGAACCG	CTGAACGTAGCAAAAGTTGGGCAGA	95504
QY	514	-----	-----	-----	-----	514
Db	95503	GGGACCGACTG	CTGGGATGACAGT	CAAGAGCTTCC	ATGGCAGGCTTGGAGATGGGGAA	95444
QY	514	-----	-----	-----	-----	514
Db	95443	GGTTTTAAGCA	AGAAAGTACCT	CGTCCCAAGAAC	CGGAACAGGATGTAGGACGCACCTTC	95384
QY	514	-----	-----	-----	-----	514
Db	95383	TTGGTGTCTT	CTGCCCTACTTT	AGCTTCTAGCT	GGGTGGCTCAGATGTTCTTGGAGTC	95324
QY	514	-----	-----	-----	-----	514
Db	95323	TCCCTGACTC	AGAGTCTGTGCCT	ATATCTGGTGTG	ACAGATGCGAGTCTTGAACACC	95264
QY	514	-----	-----	-----	-----	514
Db	95263	TCTGTCCCA	CAGACGAGGCGT	CTGAGAGGAA	CTCGTGGGTCTGGCTAGTTCTCTGGGTA	95204
QY	515	-----	-----	-----	-----	520
Db	95203	GTTATTTGG	TATCTGTAGCC	CAAGCCGTGAC	CCCTCTGCTTTTACAGGAATCAGCCTCC	95144
QY	520	-----	-----	-----	-----	540
Db	95143	TCATATCCCC	CGGACGCCAT	CCCCCGGAAAG	ATCTACGAGATCTACCTCTCTGCACA	95084
QY	540	-----	-----	-----	-----	544
Db	95083	AGCCAGAG	ACGTGAGGTGT	GTGCTGCTGCT	GTGTGGGGGGGGGAGGCA	95024

QY 544 ----- 544
Db 95023 CGTCTGTCTCTTATCGCTGGAAACCAACCAATTCCTCCCGCCAGCACTGCTGTACCTG 94964
QY 544 ----- 544
Db 94963 TTCTGACCTGGCCCGAGGGAGCACAGGGGATGAGGGCACTTAAGCTCACCTGATTGC 94904
QY 544 ----- 544
Db 94903 AACTTCCTGCTTACGGTGTCAACCAACCCCACTGCAGAGTGTGTCTCTAAACCCCAACGGT 94844
QY 545 ----- ArgLeuP 547
Db 94843 CTCTAGGACTGGCCACTGACGCTTTCCTCTCCACCCATGTTTCCCACTTGAGTTGC 94784
QY 547 roLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyProGlyValL 567
Db 94783 CCTAGCTGGCTGTGACACCTGCTGAGTCTCTATCGTTAGCTGTGGGCCCCAGAGTCC 94724
QY 567 euLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerT 587
Db 94723 TGCTACCCGGCCAGTCATCTTGGCATGGACCACTGGGGGAGCCCACTCCGACAGCT 94664
QY 587 rpSerLeuArgLeuLysLysGlnSerCysGlyGlySerTirpGlu----- 601
Db 94663 GGAGCCTCGGCTCAAAAAGCAGTCTCTGTAGGGCAGCTGGGAGGTGAGTTGGGAAGACT 94604
QY 601 ----- 601
Db 94603 ACCTGAGGCCACCAACCAATCACTGGCCCCACACTGGGCATGGCCTCCCATTCCTTTG 94544
QY 601 ----- 601
Db 94543 GCAGCTATGTGGTCTCGAGTCTCCCAAACTGAGCTGAGGTGGCCCGGAGGCTCTAAGG 94484
QY 602 ----- GlnAspValLeuHisLeu 607
Db 94483 ACTCCTAAGTTGGTTAGGCACGTACCACCAAGGCCCGCCACAGAGAGTGTCTACCTT 94424
QY 608 GlyGluGluAlaProSerHisLeuTyTyCysGlnLeuGluAlaSerAlaCysTyVal 627
Db 94423 GGTGAGAGTGGCCCTCTCATCTCTACTCTGAGTGGAGGCGGGCTGTATGTC 94364
QY 628 PheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAla 647
Db 94363 TTCACCGAGCAGTAGGCCGCTTTGGCTGTGGAGAGGCCCTTCAGCGTGGCTGCCACC 94304
QY 648 LysArgLeuLysLeuLeuPheAlaProValAlaCysThrSerLeuGluTyAsnIle 667
Db 94303 AAGCGCTCAGGCTCCTTCTGTTTGGCCCTGTGGCCTGTAGTCCCTCGAGTACACATC 94244
QY 668 ArgValTyrcysLeuHisAspThrHisAspAlaLeuLys----- 680
Db 94243 CGAGTGTACTGCTGCAGCAGACCCACGATGCTCTCAA-GGTACTGCGCGGCACAGCCCC 94185
QY 680 ----- 680
Db 94184 ATTGGCGGACCCACCCCTTGGCATCCAACTGCCTTGTCTTGGGGGTGACAGGAGGG 94125
QY 681 ----- GluValValGlnLeu 685
Db 94124 TTTGCTCTACCTCGGGCTTGTGACCAACCCCGGCACCTGACGAGGTGGTGCAGCTG 94065
QY 686 GluLysGlnLeuGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSer 705
Db 94064 GAGAAGCAGCTGGGTGGACAGCTGATCCAGAGGCCCGTGTCTGTGACCTTCAAAGACAGT 94005
QY 706 TyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeu 725
Db 94004 TACCACAACCTACTGTCTGTCATCCAGAGCTGCCAGCTCCCTGTGGGAAGCAAGCTC 93945

QY 726 LeuValSerTyrGln----- 730
Db 93944 CTTGTCACTACCA-GGTGAGACCTGTAGCCCAACGACGAGCGCTCCAGGGCTGCTGG 93886
QY 731 ----- GluIleProPheTyHis 736
Db 93885 CGCTGGGCACGTTGACGCTCTCTTTGCCCTCTCCATAGGAGATCCCTTTTACCAC 93826
QY 737 IleTrpAsnGlyThrGlnArgTyrIleuHisCysThrPheThrLeuGluArgValSerPro 756
Db 93825 ATCTGGAATGCACTCAGCAGTATCTGCACTGCACCTTCAACCTGAGCGCTCAATGCC 93766
QY 757 SerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSer 776
Db 93765 AGCACCGACCTGGCTGCAAGGTGGGTGTGGCAGTGGAGGAGATGGACAGAGC 93706
QY 777 PheSerIleAsnPheAsnIleThrLys----- 785
Db 93705 TTCAACATCAACTTAAACATCACTAAGTGGCGGCACAAGATGGTGTCTCCTAGTACCAG 93646
QY 785 ----- 785
Db 93645 CCATCCAGGAGGCTCCACCTCTCGCGTAGTGTCCCAACCAACCGCCCTAACTTGA 93586
QY 785 ----- 785
Db 93585 GCTGCCCTTCTAGCTTCTTCCAAGTTTCTGGGACATAGGTGTAGGTGTATATAGGA 93526
QY 785 ----- 785
Db 93525 GCGGAGGAGGTAGAGTTGTCTGGGGCTCTGGAGCGCTTCCCAAGAGGAGCTTTGC 93466
QY 785 ----- 785
Db 93465 ACCGGCCCCAAAGATAGGACGAGCGGAACTCTGACAGACTGCAGGTTGGGAGAAG 93406
QY 785 ----- 785
Db 93405 CATTTGGGAGCAGGCACTGAGCTGAGACATAGCCATGATGTCAGCAGAGCAATAG 93346
QY 785 ----- 785
Db 93345 TGTGGGAGCAGGAGGTGTCTCCGTCACCAACTGTGGCCAGGTGACTTCATCTCTGAG 93286
QY 785 ----- 785
Db 93285 ATCTAAAGAGTAAAGCCCTGGTGTACAGGGAGAGGATGCCCTGTGCCCCCATTTGCTG 93226
QY 785 ----- 785
Db 93225 GGGTTATCTGGAAACTCCCTGCGGAGGCGGAGGGGTCCAGAGTTCTAACTGTGTACCTG 93166
QY 786 ----- AspThrArgPheAlaGluLeuAlaLeuGluSerG 798
Db 93165 TTGCAACATCCCATCATGATACAGACACAGAGTTGTCTGAAATGCTGGCTCTGAGAGTGA 93106
QY 798 uAlaGlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgG 818
Db 93105 AGGGGGGTCCCAAGCCTGGTGGGCCAGTGCCTTCAAGATCCCTTCTCTCATTCGGCA 93046
QY 818 nLysIleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAl 838
Db 93045 AAGATCATTTACGACCTGGACCCACCTGCAAGCGGGGCGCGCTGCGCAACTCTAGC 92986
QY 838 aGlnLysLeuHisLeuAspSer----- 845
Db 92985 CCAGAAACTTCACTGGACAG-GTGGGTGGAGCTAGACAGAGGGAGGCTGGTTGGGCAAG 92927
QY 845 ----- 845
Db 92926 CGAGGTAGACAGAGGAGGCGCTGTAGGCTCTGTGCCACAGCCTCTAGTGGTATTTC 92867
QY 846 ----- HisLeuSe 848

```
Db 92866 GGGCAGCTGGAGTGGAGCCAGACTACTGTACAGACTGCCCTCTTCTTTGCAGCCATCTTAG 92807
Qy 848 rPhePheAlaSerIysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHi 868
Db 92806 CTTCCTTTGGCCTCCAGCCCAAGCCCTACAGCCNTGATCTCAACCTATGGAGGGCGGCA 92747
Qy 868 sPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnProAs 888
Db 92746 CTTCCCCAACGGCAACCTCGGCCAGCTGGCCGAGCTGTGGCCGAGCTGGGCCAGCCAGA 92687
Qy 888 pAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 92686 TGCTGGCCTCTTCACCGCTGTACAGAGGCCGAGTGC 92653
```

Search completed: October 6, 2004, 21:08:28
Job time : 9114 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 18:17:48 ; Search time 100 Seconds
(without alignments)
2892.975 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRPGIWPALLGIVLAAML.....AVAGLGQPDAGLFTVSEAC 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4787	100.0	899	11	US-09-970-944-2
2	4698.5	98.2	898	12	US-09-918-779-2
3	4698.5	98.2	898	16	US-10-624-932-2
4	4545.5	95.0	898	10	US-09-933-261-5
5	4545.5	95.0	898	11	US-09-970-944-13
6	4545.5	95.0	898	14	US-10-256-702-5
7	4545.5	95.0	898	14	US-10-240-154-16
8	4379.5	91.5	842	16	US-10-311-623-1
9	2819	58.9	544	11	US-09-970-944-14
10	2789.5	58.3	557	10	US-09-933-261-6
11	2705.5	56.5	931	11	US-10-256-702-6
12	2705.5	56.5	931	12	US-09-970-944-15
13	2705.5	56.5	931	12	US-10-087-684-35
14	2705.5	56.5	931	12	US-09-972-211-121
15	2705.5	56.5	931	12	US-10-037-417-117

16	2705.5	56.5	931	12	US-10-096-625-121
17	2705.5	56.5	1010	12	US-10-218-779-35
18	2680.5	56.0	931	11	US-09-970-944-16
19	2680.5	56.0	931	12	US-09-972-211-125
20	2680.5	56.0	931	12	US-10-096-625-125
21	2677.5	55.9	931	12	US-10-037-417-120
22	2673.5	55.8	931	11	US-09-970-944-17
23	2673.5	55.8	931	12	US-10-087-684-36
24	2673.5	55.8	931	12	US-10-218-779-36
25	2673.5	55.8	931	12	US-09-972-211-122
26	2673.5	55.8	931	12	US-10-037-417-118
27	2673.5	55.8	931	12	US-10-037-417-119
28	2673.5	55.8	931	12	US-10-096-625-122
29	2497	52.2	945	12	US-10-087-684-34
30	2497	52.2	945	12	US-10-218-779-34
31	2497	52.2	945	12	US-09-972-211-124
32	2497	52.2	945	12	US-10-037-417-121
33	2497	52.2	945	12	US-10-096-625-124
34	2493	52.1	945	12	US-10-087-684-33
35	2493	52.1	945	12	US-10-218-779-33
36	2493	52.1	945	12	US-09-972-211-123
37	2493	52.1	945	12	US-10-096-625-123
38	2490.5	52.0	946	12	US-09-972-211-56
39	2490.5	52.0	946	12	US-10-096-625-56
40	2490	52.0	943	10	US-09-933-261-7
41	2490	52.0	943	14	US-10-256-702-7
42	2484	51.9	933	12	US-10-087-684-2
43	2484	51.9	933	12	US-10-087-684-4
44	2484	51.9	933	12	US-10-218-779-2
45	2484	51.9	933	12	US-10-218-779-4

ALIGNMENTS

RESULT 1

US-09-970-944-2
; Sequence 2, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
; APPLICANT: Herрман, John L
; APPLICANT: Rastelli, Luca
; APPLICANT: Shinkets, Richard A
; TITLE OF INVENTION: No. US20030204052A1 Proteins and Nucleic Acids Encoding Same and Antibodies Directed Against these Proteins
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; PRIOR FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/237,862
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 899
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-944-2

Query Match	100.0%	Score	4787	DB 11	Length	899
Best Local Similarity	100.0%	Pred. No.	0			
Matches	899	Conservative	0	Mismatches	0	Indels
					0	Gaps
						0
QY	1	MAVRPGIWPALLGIVLAAML	RGSGAQSSATV	ANPVGANPDLL	PHFLVPEFYIVKNKP	60
Db	1	MAVRPGIWPALLGIVLAAML	RGSGAQSSATV	ANPVGANPDLL	PHFLVPEFYIVKNKP	60
QY	61	VLLVCXKAVPATQIF	FKCNGEWRQVDH	VIERTDSSGSGEPTMEVR	INVSRRQKVFGL	120
Db	61	VLLVCXKAVPATQIF	FKCNGEWRQVDH	VIERTDSSGSGEPTMEVR	INVSRRQKVFGL	120
QY	121	EYWCQCVANSSGTTKSQ	KAYIRIARLKNF	QOEPLAKESVLEOGIVL	PCRPPEGTPPAE	180
Db	121	EYWCQCVANSSGTTKSQ	KAYIRIARLKNF	QOEPLAKESVLEOGIVL	PCRPPEGTPPAE	180

```

181 VEWLNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIIVARRSASAAVIVY 240
181 VEWLNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIIVARRSASAAVIVY 240
241 VNGWSTWTSWVSCASCGRWKRSCTNPAPLNGAFCEGONVHRTVSSLLVSDVG 300
241 VNGWSTWTSWVSCASCGRWKRSCTNPAPLNGAFCEGONVHRTVSSLLVSDVG 300
301 SWPWSKWSACGLDCTHWRSECDPAPRNGEGECQGTDLDRNCTSDLCVHSASGPEDV 360
301 SWPWSKWSACGLDCTHWRSECDPAPRNGEGECQGTDLDRNCTSDLCVHSASGPEDV 360
361 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSQFQVSIKPSKADNPHL 420
361 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSQFQVSIKPSKADNPHL 420
421 LTIQPDLSLTITTYOGSLCPRODGPSPKFQLTNGHLLSPLGGGRHTLHSSPTSABEFVS 480
421 LTIQPDLSLTITTYOGSLCPRODGPSPKFQLTNGHLLSPLGGGRHTLHSSPTSABEFVS 480
481 RLSTQNYFRSLPRGTSNNMTYGTFFNLGRLMIPNTGISLLIPDAIPRGKIYEIYLTJHK 540
481 RLSTQNYFRSLPRGTSNNMTYGTFFNLGRLMIPNTGISLLIPDAIPRGKIYEIYLTJHK 540
541 PEDVRLPLAGCQTLTLLSPVSCGPPGVLLTRPVILAMDHCCGPPSPDSRLRLKXKCSGSW 600
541 PEDVRLPLAGCQTLTLLSPVSCGPPGVLLTRPVILAMDHCCGPPSPDSRLRLKXKCSGSW 600
601 EQDVHLHGEAPSHLYYCOLEASACYPVTEQGRFALVGEALSVAARLKLILLFAPVAC 660
601 EQDVHLHGEAPSHLYYCOLEASACYPVTEQGRFALVGEALSVAARLKLILLFAPVAC 660
661 TSLEYNIRVYCLHDTHDALKEVWOLEKQGLQLOEPRVLHFQKDSYHNLRLSIHDPSSL 720
661 TSLEYNIRVYCLHDTHDALKEVWOLEKQGLQLOEPRVLHFQKDSYHNLRLSIHDPSSL 720
721 WKSLLYSYQIPIFYHINWNGTQRYLHCTFTLERYSPSTSDLACKLWVQVGDGQSPSIN 780
721 WKSLLYSYQIPIFYHINWNGTQRYLHCTFTLERYSPSTSDLACKLWVQVGDGQSPSIN 780
781 FNITKDFRFAELLALAESEAGVPALVGPFAKIPFLIRQKIISLDPDPCRRGADWRTLAQK 840
781 FNITKDFRFAELLALAESEAGVPALVGPFAKIPFLIRQKIISLDPDPCRRGADWRTLAQK 840
841 LHLDSHLSFFASKSPPTAMILNLEARHFPNGNLSQAAAAVAGLQPDAGLFTVSEABEC 899
841 LHLDSHLSFFASKSPPTAMILNLEARHFPNGNLSQAAAAVAGLQPDAGLFTVSEABEC 899

```

RESULT 2

```

US-09-918-779-2
; Sequence 2, Application US/09918779
; Publication No. US20030064369A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Rastelli, Luca
; APPLICANT: Spaderna, Steven
; APPLICANT: Shimkets, Richard
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Spytek, Kimberly
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Gusev, Vladimir
; APPLICANT: Grosse, William
; APPLICANT: Alsbrook, John
; APPLICANT: Legley, Denise
; APPLICANT: Burgess, Catherine
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David

```

```

; APPLICANT: Smithson, Glennda
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-074 US
; CURRENT APPLICATION NUMBER: US/09/918,779
; CURRENT FILING DATE: 2001-07-30
; PRIOR FILING DATE: 2001-07-30
; PRIOR FILING DATE: 2000-07-28
; PRIOR FILING DATE: 2000-07-28
; PRIOR FILING DATE: 2000-08-04
; PRIOR FILING DATE: 2000-08-04
; PRIOR FILING DATE: 2000-08-08
; PRIOR FILING DATE: 2000-08-08
; PRIOR FILING DATE: 2000-08-08
; PRIOR FILING DATE: 2000-08-14
; PRIOR FILING DATE: 2000-08-14
; PRIOR FILING DATE: 2000-08-15
; PRIOR FILING DATE: 2000-08-15
; PRIOR FILING DATE: 2000-08-15
; PRIOR FILING DATE: 2000-08-16
; PRIOR FILING DATE: 2000-08-16
; PRIOR FILING DATE: 2001-02-01
; PRIOR FILING DATE: 2001-02-01
; PRIOR FILING DATE: 2001-04-05
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-918-779-2

```

Query Match 98.2%; Score 4698.5; DB 12; Length 898;

Best Local Similarity 98.7%; Pred. No. 0;

Matches 888; Conservative 2; Mismatches 7; Indels 3; Gaps 3;

```

Qy 1 MAVRPGIWPALLGIVLAALIRGSGAQSATVANVPVPGANPDLLPHFLVEPDEVIVKPK 60
Db 1 MAVRPGIWPALLGIVLAALIRGSGAQSATVANVPVPGANPDLLPHFLVEPDEVIVKPK 60
Qy 61 VLLVCKAVPATQIFPKCNGEVRQVDHVIERTDGSSEPTMEVRINVSQOQKVFGL 120
Db 61 VLLVCKAVPATQIFPKCNGEVRQVDHVIERTDGSSEPTMEVRINVSQOQKVFGL 120
Qy 121 EYWCQCVAMSSSGTTKQKAYIRIARLRKNEQEPPLAKEVSLQGIPLPCRPPEGIPPAE 180
Db 121 EYWCQCVAMSSSGTTKQKAYIRIARLRKNEQEPPLAKEVSLQGIPLPCRPPEGIPPAE 180
Qy 181 VEWLNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIIVARRSASAAVIVY 240
Db 181 VEWLNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIIVARRSASAAVIVY 240
Qy 241 VNGWSTWTSWVSCASCGRWKRSCTNPAPLNGAFCEGONVHRTVSSLLVSDVG 300
Db 241 VNGWSTWTSWVSCASCGRWKRSCTNPAPLNGAFCEGONVHRTVSSLLVSDVG 300
Qy 301 SWPWSKWSACGLDCTHWRSECDPAPRNGEGECQGTDLDRNCTSDLCVHSASGPEDV 360
Db 301 SWPWSKWSACGLDCTHWRSECDPAPRNGEGECQGTDLDRNCTSDLCVHSASGPEDV 360
Qy 361 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSQFQVSIKPSKADNPHL 420
Db 361 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSQFQVSIKPSKADNPHL 420
Qy 421 LTIQPDLSLTITTYOGSLCPRODGPSPKFQLTNGHLLSPLGGGRHTLHSSPTSABEFV 479
Db 421 LTIQPDLSLTITTYOGSLCPRODGPSPKFQLTNGHLLSPLGGGRHTLHSSPTSABEFV 479

```


QY 840 KLHLDHLSFFASKPSPTAMILNWEARHPNGLSOLAAAVAGLGPDAGLFTVSEAE 899
DB 839 KLHLDHLSFFASKPSPTAMILNWEARHPNGLSOLAAAVAGLGPDAGLFTVSEAE 898

RESULT 4
US-09-933-261-5
; Sequence 5, Application US/09933261
; Publication No. US20030040046A1
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsey
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261
; FILING DATE: 20-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/808,982
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030040046A1 Relevant
; TOPOLOGY: No. US20030040046A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-933-261-5

Query Match 95.0%; Score 4545.5; DB 10; Length 898;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

QY 1 MAVRPLMPALLGIVLAALRGSGAQSSATVANPVGANPDLLPHFLVEPDVYIVKNKP 60
DB 1 MAVRPLMPVLLGIVLAALRGSGAQSSATVANPVGANPDLLPHFLVEPDVYIVKNKP 60

QY 61 VLLVCKAVPATQIFPKNGEVRVDVHIERSTGSGEPTMEVRINVSROQVEKVGLE 120
DB 61 VLLVCKAVPATQIFPKNGEVRVDVHIERSTGSGEPTMEVRINVSROQVEKVGLE 120

QY 121 EYWCOCVWSSGGTTKSKAVIRIARLRNFEQEPFLAKEVSLQGIPLCPRPPEGIPPAE 180
DB 121 EYWCOCVWSSGGTTKSKAVIRIARLRNFEQEPFLAKEVSLQGIPLCPRPPEGIPPAE 180

QY 181 VEWLNRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVY 240
DB 181 VEWLNRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVY 240

QY 241 VNGGWTWTWSEVCSASCGRWQKRSCCTNPAPLNGAFCEQNVHRTVSSLLVSDG 300
DB 241 VNGGWTWTWSEVCSASCGRWQKRSCCTNPAPLNGAFCEQNV-QKTACATLCPVDG 299

QY 301 SWSPSKWSACGLDCTHWSRECSDPAPRNGBECOGTDLTNRCTSDLCVHSGASGEDV 360
DB 300 SWSWSKWSACGLDCTHWSRECSDPAPRNGBECRGAULTNRCTSDLCVHSGASGEDV 359

QY 361 ALYVGLIYAVAVCLVLLVLLVLYVCRKKEGLDSVDSSILTSQFOPVSTKPSKADNPHL 420
DB 360 ALYIGLVAVAVCLVLLVLLVLYVCRKKEGLDSVDSSILTSQFOPVSTKPSKADNPHL 419

QY 421 LTIQPDLS-TTTTVOGSLCPRODGPSPKFOLITNGHLLSPGLGGGRHTLHSSPSSEAEFV 479
DB 420 LTIQPDLS-TTTTVOGSLCPRODGPSPKFOLITNGHLLSPGLGGGRHTLHSSPSSEAEFV 479

QY 480 SRLSTQNYFRSLPRGTSNMTYGTFFLGGRLMTPNTGISLILPPDAIPRGIYEIYTLH 539
DB 480 SRLSTQNYFRSLPRGTSNMTYGTFFLGGRLMTPNTGISLILPPDAIPRGIYEIYTLH 539

QY 540 KPEDVRLPLAGCOTLLSPIYSCGPPGVLLTRPVILAMDHCGEPPSPDSWSRLKKQCEGS 599
DB 540 KPEDVRLPLAGCOTLLSPIYSCGPPGVLLTRPVILAMDHCGEPPSPDSWSRLKKQCEGS 599

QY 600 WEODVLHGBEAPSHLYVCOLEASACVYFTEQLGRFALVGEALSAAAKRLKLLFAPVA 659
DB 600 WE-DVLHGBEAPSHLYVCOLEASACVYFTEQLGRFALVGEALSAAAKRLKLLFAPVA 658

QY 660 CTSLEYNIRVYCLHDTDALKEVVVQLEKQGGQIQBPVRLHFKDSVHNLRLSHDVPS 719
DB 659 CTSLEYNIRVYCLHDTDALKEVVVQLEKQGGQIQBPVRLHFKDSVHNLRLSHDVPS 718

QY 720 LWSKLLVSYQETPFYHINWGTQRYLHCTETLERSVSTSDLACKLVWVQVGGQGSFI 779
DB 719 LWSKLLVSYQETPFYHINWGTQRYLHCTETLERSVSTSDLACKLVWVQVGGQGSFI 778

QY 780 NFNTKDTFAELIALALESEAGVPALVGPSAFKIPFLIRQKIISLDPCCRGADWRTLAQ 839
DB 779 NFNTKDTFAELIALALESEAGVPALVGPSAFKIPFLIRQKIISLDPCCRGADWRTLAQ 838

QY 840 KLHLDHLSFFASKPSPTAMILNWEARHPNGLSOLAAAVAGLGPDAGLFTVSEAE 899
DB 839 KLHLDHLSFFASKPSPTAMILNWEARHPNGLSOLAAAVAGLGPDAGLFTVSEAE 898

RESULT 5
US-09-970-944-13
; Sequence 13, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
; APPLICANT: Herrman, John L
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same and
; TITLE OF INVENTION: Antibodies Directed Against these Proteins
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/237,862
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 13
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-970-944-13

Query Match 95.0%; Score 4545.5; DB 11; Length 898;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

QY 1 MAVRPLMPALLGIVLAALRGSGAQSSATVANPVGANPDLLPHFLVEPDVYIVKNKP 60

```
Db 1 MAVRPGMLPVLLGIVLAAMLRGSGAQSATVANPVGANPDLLPHFLVEPEDVYIVKNKP 60
QY 61 VLLVCKAVPATQIFPKCNQWVRQVDHVIERTDSSGSEPTMEVRINVSROQVEKVFGL 120
Db 61 VLLVCKAVPATQIFPKCNQWVRQVDHVIERTDSSGSEPTMEVRINVSROQVEKVFGL 120
QY 121 EYWCQCVAWSSSGTTKSKAYIRIARLNKNEQEPLEKEVLEQIGVLPCEPPEGIPPAE 180
Db 121 EYWCQCVAWSSSGTTKSKAYIRIARLNKNEQEPLEKEVLEQIGVLPCEPPEGIPPAE 180
QY 181 VEWLNRNEDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIV 240
Db 181 VEWLNRNEDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIV 240
QY 241 VNGGWSWTWESVCSASCGRGWQKRSCTNPAPNGGECQGTDLTRNCTSDLCVHSASGPE 300
Db 241 VNGGWSWTWESVCSASCGRGWQKRSCTNPAPNGGECQGTDLTRNCTSDLCVHSASGPE 300
QY 301 SWSPWSKWSACGLDCTHWRSECSDPAPNGGECQGTDLTRNCTSDLCVHSASGPE 360
Db 300 SWSPWSKWSACGLDCTHWRSECSDPAPNGGECQGTDLTRNCTSDLCVHSASGPE 359
QY 361 ALYVGLTAVAVCLVLLIALLGLIYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL 420
Db 360 ALYVGLTAVAVCLVLLIALLGLIYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL 419
QY 421 LTIQPDLS-TTTTYQGSICPRQDGPSPKFQLTNGHLLSPGGRHTLHSSPTSEAEFV 479
Db 420 LTIQPDLS-TTTTYQGSICPRQDGPSPKFQLTNGHLLSPGGRHTLHSSPTSEAEFV 479
QY 480 SRLSTQNYFRSLPRGTSMNTYGTGFNGLGRLMIPNTGISLIPDPAIPRGKIYIYLT 539
Db 480 SRLSTQNYFRSLPRGTSMNTYGTGFNGLGRLMIPNTGISLIPDPAIPRGKIYIYLT 539
QY 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKXQCEGS 599
Db 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKXQCEGS 599
QY 600 WEODVHLGEPASHLYCOLEASACVFTQGLRFPALVGEALVAAAKRLKLLFAPVA 659
Db 600 WE-DVHLGEPASHLYCOLEAGACVFTQGLRFPALVGEALVAAATKRLRLFPVA 658
QY 660 CTSLEYNIRYVCLDTHDALKEVVQLEKQGGQIIOEPVRLHFKDSVHNLRSLTHDVPSS 719
Db 659 CTSLEYNIRYVCLDTHDALKEVVQLEKQGGQIIOEPVRLHFKDSVHNLRSLTHDVPSS 718
QY 720 LWSKLLVSYQEIPIFYHWNQTYRHLCTFTLERSVSTSDLACKLWVWQVGGQSFNI 779
Db 719 LWSKLLVSYQEIPIFYHWNQTYRHLCTFTLERSVSTSDLACKLWVWQVGGQSFNI 778
QY 780 NFENITKDTFAELALESEAGVPALVCPSAFKIPFLIROKIISSLDPPCRRGADWRTLAQ 839
Db 779 NFENITKDTFAELALESEAGVPALVCPSAFKIPFLIROKIISSLDPPCRRGADWRTLAQ 838
QY 840 KLHLDLHLSFFASKPSPTAMILNLWEARHPNGLSOLAAVAGLQPDAGLFTVSEAE 899
Db 839 KLHLDLHLSFFASKPSPTAMILNLWEARHPNGLSOLAAVAGLQPDAGLFTVSEAE 898

RESULT 6
US-10-256-702-5
; Sequence 5, Application US/10256702
; Publication No. US20030059859A1
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
```

```
;
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/256,702
; FILING DATE: 27-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261
; FILING DATE: 20-Aug-2001
; APPLICATION NUMBER: 08/808,982
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030059859A1 Relevant
; TOPOLOGY: No. US20030059859A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
; US-10-256-702-5
```

```
Query Match 95.0%; Score 4545.5; DB 14; Length 898;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;
```

```
QY 1 MAVRPGMLPVLLGIVLAAMLRGSGAQSATVANPVGANPDLLPHFLVEPEDVYIVKNKP 60
Db 1 MAVRPGMLPVLLGIVLAAMLRGSGAQSATVANPVGANPDLLPHFLVEPEDVYIVKNKP 60
QY 61 VLLVCKAVPATQIFPKCNQWVRQVDHVIERTDSSGSEPTMEVRINVSROQVEKVFGL 120
Db 61 VLLVCKAVPATQIFPKCNQWVRQVDHVIERTDSSGSEPTMEVRINVSROQVEKVFGL 120
QY 121 EYWCQCVAWSSSGTTKSKAYIRIARLNKNEQEPLEKEVLEQIGVLPCEPPEGIPPAE 180
Db 121 EYWCQCVAWSSSGTTKSKAYIRIARLNKNEQEPLEKEVLEQIGVLPCEPPEGIPPAE 180
QY 181 VEWLNRNEDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIV 240
Db 181 VEWLNRNEDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIV 240
QY 241 VNGGWSWTWESVCSASCGRGWQKRSCTNPAPNGGECQGTDLTRNCTSDLCVHSASGPE 300
Db 241 VNGGWSWTWESVCSASCGRGWQKRSCTNPAPNGGECQGTDLTRNCTSDLCVHSASGPE 299
QY 301 SWSPWSKWSACGLDCTHWRSECSDPAPNGGECQGTDLTRNCTSDLCVHSASGPE 360
Db 300 SWSPWSKWSACGLDCTHWRSECSDPAPNGGECQGTDLTRNCTSDLCVHSASGPE 359
QY 361 ALYVGLTAVAVCLVLLIALLGLIYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL 420
Db 360 ALYVGLTAVAVCLVLLIALLGLIYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL 419
QY 421 LTIQPDLS-TTTTYQGSICPRQDGPSPKFQLTNGHLLSPGGRHTLHSSPTSEAEFV 479
Db 420 LTIQPDLS-TTTTYQGSICPRQDGPSPKFQLTNGHLLSPGGRHTLHSSPTSEAEFV 479
```



```

QY 480 SRLSTQNYFRSLPRGTSNMTYCTFNFLGGLMIPNTGISLILPPDAIPRCKIYEIYTLH 539
DB 480 SRLSTQNYFRSLPRGTSNMTYCTFNFLGGLMIPNTGISLILPPDAIPRCKIYEIYTLH 539
QY 540 KPEDVRLPLAGCOTLLSPVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSCGS 599
DB 540 KPEDVRLPLAGCOTLLSPVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSCGS 599
QY 600 WEODVLHLGGEAPSHLYYCOLEASACVVFTEQLGRFALVGEALSVAALKLLPAPVA 659
DB 600 WE-DVLHLGGEAPSHLYYCOLEASACVVFTEQLGRFALVGEALSVAALKLLPAPVA 658
QY 660 CTSLEYNIRVYCLHDTDALKEVVQLEKQGGQIQQEPRVLHFKDSVHNLRLSLHDPVSS 719
DB 660 CTSLEYNIRVYCLHDTDALKEVVQLEKQGGQIQQEPRVLHFKDSVHNLRLSLHDPVSS 718
QY 720 LWSKLLVSYQEIIPFYHINWGTQOYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNI 779
DB 720 LWSKLLVSYQEIIPFYHINWGTQOYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNI 778
QY 780 NFNITKDTREAFELALAESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRGADWRTIAQ 839
DB 780 NFNITKDTREAFELALAESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRGADWRTIAQ 838
QY 840 KLHLDLSLFFASKPSPMTAMLNWEARHPNGLSQALAAVAGLGPDPAGLFTVSEAC 899
DB 840 KLHLDLSLFFASKPSPMTAMLNWEARHPNGLSQALAAVAGLGPDPAGLFTVSEAC 898

RESULT 7
US-10-240-154-16
; Sequence 16, Application US/10240154
; Publication No. US2003017541A1
; GENERAL INFORMATION:
; APPLICANT: Cochran et al.
; TITLE OF INVENTION: SCHIZOPHRENIA RELATED GENES
; FILE REFERENCE: CFM-P01-006
; CURRENT APPLICATION NUMBER: US/10/240,154
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: PCT/GB01/01486
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-240-154-16

Query Match 95.0%; Score 4545.5; DB 14; Length 898;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

QY 1 MAVRPLWPLALGLVLAALWLGSGAQSAQATVANPVPANPOLLPHFLVEPDEVIVKKNP 60
DB 1 MAVRPLWPLVLLGLVLAALWLGSGAQSAQATVANPVPANPOLLPHFLVEPDEVIVKKNP 60
QY 61 VLLVCXAVPATOIFPKCNGEVRQVDHVIERTDSSGSEPTMEVRINVSROQVEKVFGL 120
DB 61 VLLVCXAVPATOIFPKCNGEVRQVDHVIERTDSSGSEPTMEVRINVSROQVEKVFGL 120
QY 121 EYWCQVAVSSGTTKSKAVIRIARLNKNEOEPLAKEVSLQEGVLPCCRPPEGIPPAE 180
DB 121 EYWCQVAVSSGTTKSKAVIRIARLNKNEOEPLAKEVSLQEGVLPCCRPPEGIPPAE 180
QY 181 VEVLNEDLVPSLDPNVIITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVY 240
DB 181 VEVLNEDLVPSLDPNVIITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVY 240
QY 241 VNGHSTWTWESVCSASGRGQKRSCTNPAPLNGGAFCEGQNVHDTVSSLLVSDG 300
DB 241 VNGHSTWTWESVCSASGRGQKRSCTNPAPLNGGAFCEGQNVHDTVSSLLVSDG 299

```

```

QY 301 SMSWPKSWACGLDCTHWSRECSDPAPRNGGEECOGTDLDTRNCTSDLCVHSASPEDV 360
DB 300 SMSWPKSWACGLDCTHWSRECSDPAPRNGGEECRGADLDRNCTSDLCVHSASPEDV 359
QY 361 ALYVGLIAVAVCVLLVLLVLYVCKKGLDSDVADSSILTSFGQPVSIKPSKADNPHL 420
DB 360 ALYVGLIAVAVCVLLVLLVLYVCKKGLDSDVADSSILTSFGQPVSIKPSKADNPHL 419
QY 421 LTTQPDLS-TTTTYQSLCPRQDPSPKLTQNGHLLSLPLGGGRHTLHSSPSTSEAEFV 479
DB 420 LTTQPDLS-TTTTYQSLCPRQDPSPKLTQNGHLLSLPLGGGRHTLHSSPSTSEAEFV 479
QY 480 SRLSTQNYFRSLPRGTSNMTYCTFNFLGGLMIPNTGISLILPPDAIPRCKIYEIYTLH 539
DB 480 SRLSTQNYFRSLPRGTSNMTYCTFNFLGGLMIPNTGISLILPPDAIPRCKIYEIYTLH 539
QY 540 KPEDVRLPLAGCOTLLSPVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSCGS 599
DB 540 KPEDVRLPLAGCOTLLSPVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSCGS 599
QY 600 WEODVLHLGGEAPSHLYYCOLEASACVVFTEQLGRFALVGEALSVAALKLLPAPVA 659
DB 600 WE-DVLHLGGEAPSHLYYCOLEASACVVFTEQLGRFALVGEALSVAALKLLPAPVA 658
QY 660 CTSLEYNIRVYCLHDTDALKEVVQLEKQGGQIQQEPRVLHFKDSVHNLRLSLHDPVSS 719
DB 660 CTSLEYNIRVYCLHDTDALKEVVQLEKQGGQIQQEPRVLHFKDSVHNLRLSLHDPVSS 718
QY 720 LWSKLLVSYQEIIPFYHINWGTQOYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNI 779
DB 720 LWSKLLVSYQEIIPFYHINWGTQOYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNI 778
QY 780 NFNITKDTREAFELALAESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRGADWRTIAQ 839
DB 779 NFNITKDTREAFELALAESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRGADWRTIAQ 838
QY 840 KLHLDLSLFFASKPSPMTAMLNWEARHPNGLSQALAAVAGLGPDPAGLFTVSEAC 899
DB 839 KLHLDLSLFFASKPSPMTAMLNWEARHPNGLSQALAAVAGLGPDPAGLFTVSEAC 898

RESULT 8
US-10-311-623-1
; Sequence 1, Application US/10311623
; Publication No. US20040023244A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.
; APPLICANT: KALLICK, Deborah A.; TRIBOULEY, Catherine M.
; APPLICANT: YUB, Henry; NGUYEN, Dannel B.
; APPLICANT: TANG, Y. Tom; LAL, Preeti G.
; APPLICANT: POLICKY, Jennifer L.; AZIMZAI, Yalda
; APPLICANT: LU, Dyung Aina M.; GRAUL, Richard C.
; APPLICANT: YAO, Monique G.; BURFORD, Neil
; APPLICANT: HAFALIA, April J. A.; BAUGHN, Mariah R.
; APPLICANT: BANDMAN, Olga; ARVIZU, Chandra S.
; APPLICANT: YANG, Junming; XU, Yuming
; APPLICANT: GANDHI, Aneena R.; WARREN, Bridget A.
; APPLICANT: DING, Li; SANTANWALA, Madhusudan M.
; APPLICANT: DUGAN, Brendan M.; LU, Yan
; TITLE OF INVENTION: RECEPTORS
; FILE REFERENCE: PF-0793 USN
; CURRENT APPLICATION NUMBER: US/10/311,623
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US 01/19942
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/214,027
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/228,045
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/255,104
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program

```

; SEQ ID NO 1
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040023244A1 6052371CD1
US-10-311-623-1

Query Match
Best Local Similarity 91.5%; Score 4379.5; DB 16; Length 842;
Matches 838; Conservative 1; Mismatches 2; Indels 59; Gaps 3;
QY 1 MAVRPGLWALLGIVLAALRGSGAQQAATVANVPFGANPOLLPHFLVEPEDVYIVKNKP 60
DB 1 MAVRPGLWALLGIVLAALRGSGAQQAATVANVPFGANPOLLPHFLVEPEDVYIVKNKP 60
QY 61 VLLVCKAVPATQIFPKNGEWRQVDHVIERTSGSSGPTMEVRIIVSRQQVEKVFGL 120
DB 61 VLLVCKAVPATQIFPKNGEWRQVDHVIERTSGSSGPTMEVRIIVSRQQVEKVFGL 120
QY 121 EYWCQVAMSSGTTKQKAVIRIARLKNFEQEPPLAKEVLSLEQIVLPCRPPEGIPPAE 180
DB 121 EYWCQVAMSSGTTKQKAVIRIARLKNFEQEPPLAKEVLSLEQIVLPCRPPEGIPPAE 180
QY 181 VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
DB 181 VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
QY 241 VNGGWSTWTSVCSACGRGWKRSCTNPAPLNGAFCEGQNVHDRTVSSLLVSDG 300
DB 241 -----VDG 243
QY 301 SWSPKWSACGLDCTHWRSECDPAPRNGECQGTDLDRNCTSDLCVHSASGPEV 360
DB 244 SWSPKWSACGLDCTHWRSECDPAPRNGECQGTDLDRNCTSDLCVHTASGPEV 303
QY 361 ALVYGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSQFPVSIKPSKADNPHL 420
DB 304 ALVYGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSQFPVSIKPSKADNPHL 363
QY 421 LTTPQDLS -TTTTYQGSICPRQDGPSPKQFQTNHLLSPGLGGRRHTLHSSPTSEAEFV 479
DB 364 LTTPQDLS -TTTTYQGSICPRQDGPSPKQFQTNHLLSPGLGGRRHTLHSSPTSEAEFV 423
QY 480 SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRIMPNTGISLLIPDDAI PRGKIYIYLT 539
DB 424 SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRIMPNTGISLLIPDDAI PRGKIYIYLT 483
QY 540 KPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQCEGS 599
DB 484 KPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQCEGS 543
QY 600 WEQDVHLGGEAPSHLYYCOLEASACVVFTEQGRFALVGEALSVAIAAKRLKLLFPAP 659
DB 544 WE -DVLHLGGEAPSHLYYCOLEASACVVFTEQGRFALVGEALSVAIAAKRLKLLFPAP 602
QY 660 CTSLEYNIRVYCLDTHDALKEVVOLEKQGGQLIQEPVRLHFKDSYHNLRSLSIHDVPS 719
DB 603 CTSLEYNIRVYCLDTHDALKEVVOLEKQGGQLIQEPVRLHFKDSYHNLRSLSIHDVPS 662
QY 720 LWSKLLVSVQEIIPFYHWNQYRLHCTFTLERVSPSTSDACKLWVWQVEGDGQSF 779
DB 663 LWSKLLVSVQEIIPFYHWNQYRLHCTFTLERVSPSTSDACKLWVWQVEGDGQSF 722
QY 780 NFNITKDTFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIISSLDPPCRRGADWRTL 839
DB 723 NFNITKDTFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIISSLDPPCRRGADWRTL 782
QY 840 KLHLDLSHLSFASKPSPTAMILNWEARHPNGNLSQLAAAVAGLQPDAGLFTVSEAC 899
DB 783 KLHLDLSHLSFASKPSPTAMILNWEARHPNGNLSQLAAAVAGLQPDAGLFTVSEAC 842

RESULT 9

US-09-970-944-14
; Sequence 14, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
; APPLICANT: Herriman, John L
; APPLICANT: Rastelli, Luca
; APPLICANT: Shmukets, Richard A
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same and
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; PRIOR FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/237,862
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-944-14

Query Match 58.9%; Score 2819; DB 11; Length 544;
Best Local Similarity 99.6%; Pred. No. 4.2e-238;
Matches 540; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 359 DVALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSQFPVSIKPSKADNP 418
DB 4 DVALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSQFPVSIKPSKADNP 63
QY 419 HLLTIQPDLS -TTTTYQGSICPRQDGPSPKQFQTNHLLSPGLGGRRHTLHSSPTSEAE 477
DB 64 HLLTIQPDLS -TTTTYQGSICPRQDGPSPKQFQTNHLLSPGLGGRRHTLHSSPTSEAE 123
QY 478 FVSRSTQNYFRSLPRGTSNMTYGTFFNLGGRIMPNTGISLLIPDDAI PRGKIYIYLT 537
DB 124 FVSRSTQNYFRSLPRGTSNMTYGTFFNLGGRIMPNTGISLLIPDDAI PRGKIYIYLT 183
QY 538 LHPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQCE 597
DB 184 LHPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQCE 243
QY 598 GSWEQDVHLGGEAPSHLYYCOLEASACVVFTEQGRFALVGEALSVAIAAKRLKLLFP 657
DB 244 GSWE -DVLHLGGEAPSHLYYCOLEASACVVFTEQGRFALVGEALSVAIAAKRLKLLFP 302
QY 658 VACTSLEYNIRVYCLDTHDALKEVVOLEKQGGQLIQEPVRLHFKDSYHNLRSLSIHDV 717
DB 303 VACTSLEYNIRVYCLDTHDALKEVVOLEKQGGQLIQEPVRLHFKDSYHNLRSLSIHDV 362
QY 718 SSLWSKLLVSVQEIIPFYHWNQYRLHCTFTLERVSPSTSDACKLWVWQVEGDGQSF 777
DB 363 SSLWSKLLVSVQEIIPFYHWNQYRLHCTFTLERVSPSTSDACKLWVWQVEGDGQSF 422
QY 778 SINFNITKDTFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIISSLDPPCRRGADWRTL 837
DB 423 SINFNITKDTFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIISSLDPPCRRGADWRTL 482
QY 838 AQKHLDSHLSFASKPSPTAMILNWEARHPNGNLSQLAAAVAGLQPDAGLFTVSE 897
DB 483 AQKHLDSHLSFASKPSPTAMILNWEARHPNGNLSQLAAAVAGLQPDAGLFTVSE 542
QY 898 EC 899
DB 543 EC 544

RESULT 10

US-09-933-261-6
; Sequence 6, Application US/09933261
; Publication No. US20030040046A1

GENERAL INFORMATION:		360	XD	SYHNLXLSDHVPSSLMKSLVSYQEIFYHWNQYRHLCTFTLERSVSPSTDLA	419
APPLICANT: Tessier-Lavigne, Marc					
Leonardo, E. David					
Hink, Lindsay					
Masu, Masayuki					
Kazuko, Keino-Masu					
TITLE OF INVENTION: Netrin Receptors					
NUMBER OF SEQUENCES: 8					
CORRESPONDENCE ADDRESS:					
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP					
STREET: 268 BUSH STREET, SUITE 3200					
CITY: SAN FRANCISCO					
STATE: CALIFORNIA					
COUNTRY: USA					
ZIP: 94104					
COMPUTER READABLE FORM:					
MEDIUM TYPE: Floppy disk					
COMPUTER: IBM PC compatible					
OPERATING SYSTEM: PC-DOS/MS-DOS					
SOFTWARE: Patent in Release #1.0, Version #1.30					
CURRENT APPLICATION DATA:					
APPLICATION NUMBER: US/09/933,261					
FILING DATE: 20-Aug-2001					
CLASSIFICATION: <Unknown>					
PRIOR APPLICATION DATA:					
APPLICATION NUMBER: 08/808,982					
FILING DATE: <Unknown>					
ATTORNEY/AGENT INFORMATION:					
NAME: OSMAN, RICHARD A					
REGISTRATION NUMBER: 36,627					
REFERENCE/DOCKET NUMBER: UC96-217					
TELECOMMUNICATION INFORMATION:					
TELEPHONE: (415) 343-4341					
TELEFAX: (415) 343-4342					
INFORMATION FOR SEQ ID NO: 6:					
SEQUENCE CHARACTERISTICS:					
LENGTH: 557 amino acids					
TYPE: amino acid					
STRANDEDNESS: No. US20030040046A1 Relevant					
TOPOLOGY: No. US20030040046A1 Relevant					
MOLECULE TYPE: peptide					
SEQUENCE DESCRIPTION: SEQ ID NO: 6:					
US-09-933-261-6					
Query Match					
Best Local Similarity					
Matches 538; Conservative					
2; Mismatches 15; Indels					
3; Gaps					
3; Gaps					
QY	344	NCTSLCVHSGAGPDEVVGLIYGLIAVAVCLVLLVLLVLYCRKKEGLDSDVADSSILTS	403		
Db	1	NCTSLXVHTASGPDVAVLYGLIAVAVCLVLLVLLVLYCRKKEGLDSDVADSSILTS	60		
QY	404	GFQPVSIKPSKADNPHLLTIQPDLS-TTTTYQGSICPRQDGSPKQLTNGHLLSPLGGG	462		
Db	61	GFQPVSIKPSKADNPHLLTIQPDLS-TTTTYQGSICPRQDGSPKQLTNGHLLSPLGGG	120		
QY	463	RHTLHSSPTSEAEFEVSRSLTQNYFRSLPRGTSNMTYGTNFTLGGRLMIPNTGISLLIP	522		
Db	121	RHTLHSSPTSEAEFEVSRSLTQNYFRSLPRGTSNMTYGTNFTLGGRLMIPNTGISLLIP	180		
QY	523	PDALPRGKIYIYLTTHKPEDVRLPLAGCQTLLSPVSCGPPGVLLTPFVILAMDHCGEP	582		
Db	181	PDALPRGKIYIYLTTHKPEDVRLPLAGCQTLLSPVSCGPPGVLLTPFVILAMDHCGEP	240		
QY	583	SPDWSLRKQSCGSGWEQDVLHIGEPAPSHLYYQLEASACYVFTQLGRFALVGEAL	642		
Db	241	SPDWSLRKQSCGSGWE-DVLHIGEPAPSHLYYQLEASACYVFTQLGRFALVGEAL	299		
QY	643	SVAAAKRLKLLLPFVACTSLEYNRYVCLHDTHTALKEVVQLEKQLGGQLIQEPRVLHF	702		
Db	300	SVAAAKRLKLLLPFVACTSLEYNRYVCLHDTHTALKEVVQLEKQLGGQLIQEPRVLHL	359		
QY	703	KDSYHNLRLSIHDPVSSLMKSLVSYQEIFYHWNQYRHLCTFTLERSVSPSTDLA	762		

QY 404 GFQVSIKPSKADNPHTLTIQPDLS-FTTYYQSLCPQDQSPKFKOLTNGHLLSPGLGG 462
 Db 61 GFQVSIKPSKADNPHTLTIQPDLSFTTYYQSLCPQDQSPKFKOLTNGHLLSPGLGG 120
 QY 463 RHTLHSSPTSEAEFEVSRSLSTQNYFRSLPRGTSNMTYGTNFGLLGRMLIPNTGISLLIP 522
 Db 121 RHTLHSSPTSEAEFEVSRSLSTQNYFRSLPRGTSNMTYGTNFGLLGRMLIPNTGISLLIP 180
 QY 523 PDAIPRGKIYEIYTLTKHPEDVRLPLAGCQTLSPISVSCGPPGVLLTRPVLAMDHGCEP 582
 Db 181 PDAIPRGKIYEIYTLTKHPEDVRLPLAGCQTLSPISVSCGPPGVLLTRPVLAMDHGCEP 240
 QY 583 SPDSWSLRLKQSCGSEWQDVHLHGBEAPSHLYYCOLEASACVVFTEQLGRFALVGEAL 642
 Db 241 SPDSWSLRLKQSCGSEWQDVHLHGBEAPSHLYYCOLEASACVVFTEQLGRFALVGEAL 299
 QY 643 SVAARLKLKLLFAPVACTSLEYNIRVYCLDTHDALKEVVOLEKQLGGQLIQEPRVLIHF 702
 Db 300 SVAARLKLKLLFAPVACTSLEYNIRVYCLDTHDALKEVVOLEKQLGGQLIQEPRVLIHF 359
 QY 703 KDSVHNLRLSIHDVPSLSLWKSLLVSYQETPFVHIWNGTORYLHCTFTLERSVSPSDDL 762
 Db 360 XDSVHNLRLSIHDVPSLSLWKSLLVSYQETPFVHIWNGTORYLHCTFTLERSVSPSDDL 419
 QY 763 CKLWVQVEGDGQSFNFINITKOTRFAELLALASEAGVPAVGPFAFKIPFLIROKILIS 822
 Db 420 CKLWVQVEGDGQSFNFINITKOTRFAELLALASEAGVPAVGPFAFKIPFLIROKILIS 479
 QY 823 SLDPCCRRGADWRTLAKQLHLDLHSHLFFASKPSFTAMILNLWEARHPNGLSQAALAAVA 882
 Db 480 SLDPCCRRGADWRTLAKQLHLDLHSHLFFASKPSFTAMILNLWEARHPNGLSQAALAAVA 539
 QY 883 GLQCPDAGLFT-VSEAE 899
 Db 540 GTXPAGRWLLSQCSEAE 557

RESULT 12

US-09-970-944-15
 ; Sequence 15, Application US/09970944
 ; Publication No. US20030204052A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Heriman, John L
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Shinkets, Richard A
 ; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same an
 ; TITLE OF INVENTION: Antibodies Directed Against these Proteins
 ; FILE REFERENCE: 21402-138
 ; CURRENT APPLICATION NUMBER: US/09/970,944
 ; CURRENT FILING DATE: 2002-05-02
 ; PRIOR APPLICATION NUMBER: 60/237,862
 ; PRIOR FILING DATE: 2000-10-04
 ; NUMBER OF SEQ ID NOS: 62
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 15
 ; LENGTH: 931
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-09-970-944-15

Query Match 56.5%; Score 2705.5; DB 11; Length 931;
 Best Local Similarity 56.4%; Pred. No. 9e-228;
 Matches 515; Conservative 153; Mismatches 214; Indels 31; Gaps 12;
 QY 9 PALLGIVLAWLGRGSAQQA---TVANPVPGANPDLLPHFLVEPEDVIVTKNKPVLVLC 65
 Db 26 PAL--ALLSASGTGSAQDDFFPHELPEPDPPEPLPHFLIEPEAVIVTKNKPVLNLYC 83
 QY 66 KAVPATQIFPKCNQWVRQDVHVIERSSTGSSGEPTMEVRINVSROQVKVFGLEIYWCQ 125
 Db 84 KASPATQIVFKCNSEWVHOKDVDRVDETSGLIVREVSIBISROQVBEFGPDYWCQ 143

QY 126 CVAWSSGTTKSKAYIRIARLKNKPOEPLAKEVSLQGIIVLPCRPEGIPPAEVEWLR 185
 Db 144 CVAWSSAGITKSKAYIRIARLKTFEQEPGLKEVSLQGEVLLQCRPEGIPPAEVEWLR 203
 QY 186 NEDLVDPSPDPNYYITREHSLVVRQARLADTANYTCVAKNIVARRRRSASAAVIVYVNGW 245
 Db 204 NEDIIDPAEDRNFYIIDHNLIIKQARLSTANYTCVAKNIVAKRKTATVIVYVNGW 263
 QY 246 STWTEWSVCASCGRWQKRSRSTAPALNGGAFCEGQVHRTVSSLLSVSDGSPW 305
 Db 264 STWTEWSVCNRCRGYQKRTCTNPAPLNGGAFCEGQSV-QKIACTTCLPVDGRWTSW 322
 QY 306 SKWSACGLDCTHWRSECDPAPNGEGECQGTDLTRNCTSDLCVHSASGPDVALVYG 365
 Db 323 SKWSTGTECTHWRRECTAPAPNGGKCDGLVLOSCKNCTDGLCMAQAPSDDDVALVYG 382
 QY 366 L-TAVAVCLVLLLVLLVYCRKKEGSDVADSSILTSFGFQVSIKPSKADNPHTLTIQ 424
 Db 383 IVIATVCLAITVVVALFVYRKNHRDPESDIIDSSALNGGFQVSIKPSKADNPHTLTIQ 439
 QY 425 PDL-STYTTTYQSLCPQDQSPKFKOLTNGHLLSPGLGGHRLTHSS-PTSEAEFEV 479
 Db 440 PDLTSAAMYRGVYALHD-VSDKIPTMTNSPILDPFLNFKIKVYNSGAVTPQODLAEFS 498
 QY 480 SRLS---TONYF-----RSLPRGT--SNMTYGTENFLGGLMTPTNGISLLIPDDA 525
 Db 499 SKLSPQMTGSLLENEALNKNQSLARQTDPSCTAFGTNSLGGHLLIPNSGVSLIPAGA 558
 QY 526 IPRGKIYEIYTLTKHPEDVRLPLAGCQTLSPISVSCGPPGVLLTRPVLAMDHGCRPSPD 585
 Db 559 IPQGRVYEMVTVHRKENMRPMDSTQLLTVPVSCGPPGALLTRPVILTHHCADPSTE 618
 QY 586 SWSLRLKQSCGSEWQDVHLHGBEAPSHLYYCOLEASACVVFTEQLGRFALVGEAL 645
 Db 619 DMKIQLKNOAVQOWE-DVVVVGEEFTTFCYIQLDAEACHILTENLSTVALVQSTTKA 677
 QY 646 AAKRLKLLFAPVACTSLEYNIRVYCLDTHDALKEVVOLEKQLGGQLIQEPRVLIHF 705
 Db 678 AAKRLKLLFAPVACTSLEYNIRVYCLDTHDALKEVVOLEKQLGGQLIQEPRVLIHF 737
 QY 706 YHNLRLSIHDVPSLSLWKSLLVSYQETPFVHIWNGTORYLHCTFTLERSVSPSDDLACKL 765
 Db 738 IHNLRSLIHDIAHSLWKSLLVSYQETPFVHIWNGTORYLHCTFTLERSVSPSDDLACKL 797
 QY 766 WYVQVEGDGQSFNFINITKOTRFAELLALASEAGVPAVGPFAFKIPFLIROKILIS 825
 Db 798 CVRQVEGEGQIFQNLCTVSEETGIDPLDIPASTITTTVTGSAFSLPLIRQKLCSSLD 857
 QY 826 PPCRRGADWRTLAKQLHLDLHSHLFFASKPSFTAMILNLWEARHPNGLSQAALAAVAGL 885
 Db 858 APOTRGHWRMLAHKUNLDRYLNIFATKSSPTGVILLDLWEAQNFDPGNLSMLAAVLEEMG 917
 QY 886 QPDAGLFTVSEAE 898
 Db 918 RHETVVSLLAEQ 930

RESULT 13

US-10-087-684-35
 ; Sequence 35, Application US/10087684
 ; Publication No. US20040029116A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Edinger, Shlomit R.
 ; APPLICANT: MacDougall, John R.
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Stone, David J.
 ; APPLICANT: Grosse, William M.
 ; APPLICANT: Lepley, Denise M.
 ; APPLICANT: Rieger, Daniel K.
 ; APPLICANT: Burgess, Catharine E.
 ; APPLICANT: Casman, Stacie, J.
 ; APPLICANT: Spytek, Kimberly A.


```

; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 121
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-972-211-121

Query Match          56.5%; Score 2705.5; DB 12; Length 931;
Best Local Similarity 56.4%; Pred. No. 9e-228;
Matches 515; Conservative 153; Mismatches 214; Indels 31; Gaps 12;

QY      9   PALLGIVLAWLGRSGGAQOSA---TVANVPFGANPDLPLPHLVEPEDVVIYKKNPVLLVC 65
Db      26   PAL--ALLSASGTCSAAQDDEFFHELPTFPSPDPPELPHELIBEEAYIYKNGKPNLYC 83

QY      66   KAVPATQIQFKKCGEWWRVQVDHVIERTSDGSSGEPTMEVRINVSROQVKFGLSEYWCQ 125
Db      84   KASPATQIQFKCNSEWHVQKHVVDERVDETSLGIIVESEISIRQOVELFGEDYWCQ 143

QY     126   CVAWSSSGTTKSOKAYIRIARLRKNFRQEPLEAKEVLSLEQGVLPCRPPEGIPPAEVEMLR 195
Db     144   CVAWSSAGITTSKAYVRIAXLKTFQEPLGKEVLSLEQEVLLQCRPEGPVPAEVEMLK 203

QY     186   NEDLVPSLDPNVIYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVYYVNGGW 245
Db     204   NEDIIDPAEDRNFIYITDHNLIIKOARLSDTANYTCVAKNIVAKRKSTTATVIYVNGGW 263

QY     246   STWTWSVCSASCGRGWKESRSCTNPAPLNGAFCEGQNVHDTVSSLVSVGDGSWPW 305
Db     264   SIWTWSVNCNSRCRGYQKKTRICTNPAPLNGAFCEGQSV-QKIACITTLCPVDGRTSW 322

QY     306   SKWSACGLDTHWRSECDSDPAPRNGEGECQGTDLDRNCTSLCVHSASGPDVALYVG 365
Db     323   SKWSTCGTECHWRRECTAPAPKNGKDCDGLVLQSKNCTDGLCWQAAPDSDDVALYVG 382

QY     366   L-IYAVAVCLVLLLILLYVYCKKEGLSDVDASSILTSGFQPVSIKPKADNPHELLTIQ 424
Db     383   IYIAVTVCUAIITVVVAFVYKNNHRDFESDIIDSSALNGGFQPVNIKAAQRD---LLAVP 439

QY     425   PDL-STTTTTYQGSICPRQDGSPKFOLTNHLLSLPIGGGRHTLHHSS-----PTSEAEFEV 479
Db     440   PDLTSAAMYRGFVVALHD-VSDKIPTNTSNPIDLPFLNKIKVYNSSGAVTPOODDAEF 498

QY     480   SRLS----TQNYF-----RSLPRGT--SNMTYGTNFNGLRLMINTGTSILLIPPPA 525
Db     499   SKLSPQMTQSLLENEALNKLNQSLAQTDPSCTAFGTFSNLGHLIIFNPSGVSLIIPAGA 558

QY     526   IPRGKIYEYLTLHKEDVRLPIAGCQTLLSPITSVCGPGVLLTRPVILAMDHCGPSPD 585
Db     559   IFQGRYENYVTVHRRKNMRPFMEDSQTLITPVWSCGPPGALLTRVILTLHHCADPSTE 618

QY     586   SWSRLRKKSCSGSEMQDVLIHLGEEAPSLLYYCOLLEASACYVFTEQLGRFALVGEALSVA 645

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 18:16:33 ; Search time 27 Seconds
(without alignments)
1718.954 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRPGMLPALLGIVLAWL.....AVAGLGQPDAGLFTVSEAC 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
 - 2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
 - 3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
 - 4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
 - 5: /cgn2_6/ptodata/2/iaa/PCPUS COMB.pep.*
 - 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4545.5	95.0	898	2	US-08-808-982-5
2	4545.5	95.0	898	3	US-09-306-902A-5
3	2789.5	58.3	557	2	US-08-808-982-6
4	2789.5	58.3	557	3	US-09-306-902A-6
5	2490.5	52.0	943	2	US-08-808-982-7
6	2490.5	52.0	943	3	US-09-306-902A-7
7	294	6.1	102	2	US-08-808-982-8
8	294	6.1	102	3	US-09-306-902A-8
9	273.5	5.7	1172	1	US-08-313-288B-19
10	249	5.2	441	3	US-08-985-526-3
11	246	5.1	239	5	PCT-US93-01652-1
12	246	5.1	1170	1	US-08-313-288B-20
13	235.5	4.9	218	3	US-08-985-526-1
14	234	4.9	1395	3	US-09-540-245A-15
15	216.5	4.5	1651	3	US-09-540-245A-18
16	208.5	4.4	469	1	US-08-313-288B-15
17	205.5	4.3	1381	3	US-09-540-245A-16
18	198.5	4.1	788	2	US-08-918-914-4
19	194.5	4.1	1069	4	US-09-877-730-2
20	194.5	4.1	1150	4	US-09-877-730-8
21	189	3.9	904	4	US-09-877-730-6
22	189	3.9	985	4	US-09-877-730-10
23	188	3.9	1297	3	US-09-540-245A-17
24	186	3.9	380	4	US-09-877-730-4
25	186	3.9	1266	4	US-08-506-296B-4
26	168	3.5	481	4	US-09-130-491-8
27	167.5	3.5	905	4	US-09-369-364A-9

28	166	3.5	837	4	US-09-122-126B-2	Sequence 2, Appli
29	166	3.5	837	4	US-09-634-286A-2	Sequence 2, Appli
30	160.5	3.4	949	4	US-09-568-559-2	Sequence 2, Appli
31	160.5	3.4	967	4	US-09-130-491-2	Sequence 2, Appli
32	159.5	3.3	727	4	US-09-445-023A-1	Sequence 1, Appli
33	159.5	3.3	1224	4	US-09-930-872-4	Sequence 4, Appli
34	158	3.3	608	4	US-09-130-491-13	Sequence 13, Appli
35	156	3.3	997	4	US-09-369-364A-7	Sequence 7, Appli
36	153.5	3.2	1059	4	US-09-907-794A-290	Sequence 290, App
37	153.5	3.2	1059	4	US-09-905-125A-290	Sequence 290, App
38	153.5	3.2	1059	4	US-09-902-775A-290	Sequence 290, App
39	153.5	3.2	1119	4	US-09-907-794A-294	Sequence 294, App
40	153.5	3.2	1119	4	US-09-905-125A-294	Sequence 294, App
41	153.5	3.2	1119	4	US-09-902-775A-294	Sequence 294, App
42	153	3.2	1081	4	US-09-369-364A-17	Sequence 17, Appl
43	152.5	3.2	757	4	US-09-963-791-24	Sequence 24, Appl
44	152.5	3.2	908	4	US-09-963-791-2	Sequence 2, Appli
45	152.5	3.2	1911	1	US-08-348-006B-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-808-982-5
; Sequence 5, Application US/08808982
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, B. David
; APPLICANT: Hink, Lindsey
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-808-982-5

Query Match 95.0%; Score 4545.5; DB 2; Length 898;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

QY 1 MAVRPGMLPALLGIVLAWLRCGSAQOQSVATVPVGANPDLPLPHLVPEDYIVKNKP 60
|||||

Db 1 MAVRPGMLPVLGIVLAALRGSGAQOSATVANPVGANPDLLPHFLVEPDEVYIVKNKP 60
QY 61 VLLVCKAVPATQIPFFKCNQEWVRQVDHVIERSITDSSGSEPTMEVRINVSQOQEVKVFGL 120
Db 61 VLLVCKAVPATQIPFFKCNQEWVRQVDHVIERSITDSSGSEPTMEVRINVSQOQEVKVFGL 120
QY 121 EYWCQCVAMSSSGTTKQKAYIRIARLKNPEQPLAKEVSLGQIVLPCRPPGIPPAE 180
Db 121 EYWCQCVAMSSSGTTKQKAYIRIARLKNPEQPLAKEVSLGQIVLPCRPPGIPPAE 180
QY 181 VEWLNRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSAAVIVY 240
Db 181 VEWLNRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSAAVIVY 240
QY 241 VNGGWSWTWMSVCSACGRGWQKRSCTNPAPNGGAFCEGQNVHRTVSSLLVSDG 300
Db 241 VNGGWSWTWMSVCSACGRGWQKRSCTNPAPNGGAFCEGQNVHRTVSSLLVSDG 300
QY 301 SWSWKSWSACGLDCTHWRSECDPAPRNGEGECQGTDLDRNCTSDLCVHSAAGPDEV 360
Db 300 SWSWKSWSACGLDCTHWRSECDPAPRNGEGECQGTDLDRNCTSDLCVHSAAGPDEV 360
QY 361 ALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSFGFQVSIKPSKADNPHL 420
Db 360 ALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSFGFQVSIKPSKADNPHL 420
QY 421 LTIQPDLS-TTTTYQGSICPRQDQSPKQFQTNHGLLSPLGGGRHTLHSSPTSEAEFV 479
Db 420 LTIQPDLS-TTTTYQGSICPRQDQSPKQFQTNHGLLSPLGGGRHTLHSSPTSEAEFV 479
QY 480 SRLSTQNYFRSLPRGTSNMAYGTNFGGRLMIPNTGISLLIPDPAIPRGKIYIYTLH 539
Db 480 SRLSTQNYFRSLPRGTSNMAYGTNFGGRLMIPNTGISLLIPDPAIPRGKIYIYTLH 539
QY 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPSPDSMSLRLKQSCGS 599
Db 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPSPDSMSLRLKQSCGS 599
QY 600 WEQVILHGBAPSHLYCQLEASACVYFTQOLGRFALVGEALSVAARKLILLFAPVA 659
Db 600 WE-DVLHGEESPSHLYCQLEAGACVYFTQOLGRFALVGEALSVAARKLILLFAPVA 658
QY 660 CTSLEYNIRVCLDTHDALKEVVLQKOLGQLIQEPVRLHFKDSYHNLRLSLHDVPSS 719
Db 659 CTSLEYNIRVCLDTHDALKEVVLQKOLGQLIQEPVRLHFKDSYHNLRLSLHDVPSS 718
QY 720 LWSKLLVSYQEIPIFYHIWNGTQRYLHCTFLERVPSTSDLACKLWVWQVEGQGSFI 779
Db 719 LWSKLLVSYQEIPIFYHIWNGTQRYLHCTFLERVPSTSDLACKLWVWQVEGQGSFI 778
QY 780 NFNTKTQTRFAELLALAESEAGVPALVGSAPKIPFLIRQKILSLDPPCRGADWRTIAQ 839
Db 779 NFNTKTQTRFAELLALAESEAGVPALVGSAPKIPFLIRQKILSLDPPCRGADWRTIAQ 838
QY 840 KLHLDLSLFSFASKPSPTAMILNLWEARHFPNGNLSQAAAVALGQPDAGLFTVSEAC 899
Db 839 KLHLDLSLFSFASKPSPTAMILNLWEARHFPNGNLSQAAAVALGQPDAGLFTVSEAC 898

RESULT 2

US-09-306-902A-5

; Sequence 5, Application US/09306902A

; Patent No. 6277585

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; Leonardo, E. David

; Hink, Lindsay

; Masu, Masayuki

; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/306.902A
FILING DATE: 07-May-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC96-217
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 898 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-306-902A-5

Query Match 95.0%; Score 4545.5; DB 3; Length 898;

Best Local Similarity 94.7%; Pred. No. 0;

Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

QY 1 MAVRPGMLPVLGIVLAALRGSGAQOSATVANPVGANPDLLPHFLVEPDEVYIVKNKP 60
Db 1 MAVRPGMLPVLGIVLAALRGSGAQOSATVANPVGANPDLLPHFLVEPDEVYIVKNKP 60
QY 61 VLLVCKAVPATQIPFFKCNQEWVRQVDHVIERSITDSSGSEPTMEVRINVSQOQEVKVFGL 120
Db 61 VLLVCKAVPATQIPFFKCNQEWVRQVDHVIERSITDSSGSEPTMEVRINVSQOQEVKVFGL 120
QY 121 EYWCQCVAMSSSGTTKQKAYIRIARLKNPEQPLAKEVSLGQIVLPCRPPGIPPAE 180
Db 121 EYWCQCVAMSSSGTTKQKAYIRIARLKNPEQPLAKEVSLGQIVLPCRPPGIPPAE 180
QY 181 VEWLNRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSAAVIVY 240
Db 181 VEWLNRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSAAVIVY 240
QY 241 VNGGWSWTWMSVCSACGRGWQKRSCTNPAPNGGAFCEGQNVHRTVSSLLVSDG 300
Db 241 VNGGWSWTWMSVCSACGRGWQKRSCTNPAPNGGAFCEGQNVHRTVSSLLVSDG 300
QY 301 SWSWKSWSACGLDCTHWRSECDPAPRNGEGECQGTDLDRNCTSDLCVHSAAGPDEV 360
Db 300 SWSWKSWSACGLDCTHWRSECDPAPRNGEGECQGTDLDRNCTSDLCVHSAAGPDEV 360
QY 361 ALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSFGFQVSIKPSKADNPHL 420
Db 360 ALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSFGFQVSIKPSKADNPHL 420
QY 421 LTIQPDLS-TTTTYQGSICPRQDQSPKQFQTNHGLLSPLGGGRHTLHSSPTSEAEFV 479
Db 420 LTIQPDLS-TTTTYQGSICPRQDQSPKQFQTNHGLLSPLGGGRHTLHSSPTSEAEFV 479
QY 480 SRLSTQNYFRSLPRGTSNMAYGTNFGGRLMIPNTGISLLIPDPAIPRGKIYIYTLH 539
Db 480 SRLSTQNYFRSLPRGTSNMAYGTNFGGRLMIPNTGISLLIPDPAIPRGKIYIYTLH 539
QY 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPSPDSMSLRLKQSCGS 599
Db 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPSPDSMSLRLKQSCGS 599

Db 540 KPDVRLPLAGCQTLSPVWCGPPGVLLTRPVILAMDHCGEPPDSWSLRKQSCGS 599
QY 600 WEQDVLHGBEAPSHLYYCOLEASACVYFTQLGREALVGEALSVAAAKRLKILLFAPVA 659
Db 600 WE-DVLHGBEAPSHLYYCOLEAGACVYFTQLGREALVGEALSVAAATKRLRLLFAPVA 658
QY 660 CTSLEYNIRVYCLDTHDALKEVVQLEKQGGQIQPRVLFHFKDSYHNLRLSLHDVPSS 719
Db 659 CTSLEYNIRVYCLDTHDALKEVVQLEKQGGQIQPRVLFHFKDSYHNLRLSLHDVPSS 718
QY 720 LWSKLLVSYQEIPIFYHWNQYRHLCTFTFLRVSPSTSLACKLWVWQVGGQGSFI 779
Db 719 LWSKLLVSYQEIPIFYHWNQYRHLCTFTFLRVSPSTSLACKLWVWQVGGQGSFI 778
QY 780 NFNTTKTRFAELLALSEAGVPALVGPSPAKIPFLIRQKIISLDPPCRGADWRTIAQ 839
Db 779 NFNTTKTRFAELLALSEAGVPALVGPSPAKIPFLIRQKIISLDPPCRGADWRTIAQ 838
QY 840 KLHLDLSHLSFFASKPSPMTAMILNLWEARHPNGNLSQLAAVAGLGQPDAGLFTVSEAC 899
Db 839 KLHLDLSHLSFFASKPSPMTAMILNLWEARHPNGNLSQLAAVAGLGQPDAGLFTVSEAC 898

RESULT 3
US-08-808-982-6
; Sequence 6, Application US/08808982
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-808-982-6

Query Match 58.3%; Score 2789.5; DB 2; Length 557;
Best Local Similarity 96.4%; Pred. No. 2.7e-256;
Matches 538; Conservative 15; Indels 3; Gaps 3;
QY 344 NCTSDLCVHSAGPEVALYVGLIAVAVCLVLLVLLVYCRKKEGLSDSDVADSSILTS 403
||||| :|:|||||

Db 1 NCTSDLVHTASGPEVALYVGLIAVAVCLVLLVLLVYCRKKEGLSDSDVADSSILTS 60
QY 404 GFQVSIKPSKADNPHLLTIQPDLS-TTTTYQGSCLCPQDQSPKFKQLTNGHLLSPGGG 462
Db 61 GFQVSIKPSKADNPHLLTIQPDLS-TTTTYQGSCLCPQDQSPKFKQLTNGHLLSPGGG 120
QY 463 RHTLHHSSPTSEAEFVSRSLSTQNYFSLPRGTSNMVYGTNFGGRLMIENGTGILLIP 522
Db 121 RHTLHHSSPTSEAEFVSRSLSTQNYFSLPRGTSNMVYGTNFGGRLMIENGTGILLIP 180
QY 523 PDAIPRKIYEIYLTLLHKPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEP 582
Db 181 PDAIPRKIYEIYLTLLHKPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEP 240
QY 583 SPDSWSLRKQSCGSEWODVHLHGBEAPSHLYYCOLEASACVYFTQLGREALVGEAL 642
Db 241 SPDSWSLRKQSCGSEW-DVHLHGBEAPSHLYYCOLEASACVYFTQLGREALVGEAL 299
QY 643 SVAAAKRLKLLLPAPVACTSLEYNIRVYCLDTHDALKEVVQLEKQGGQIQPRVLFH 702
Db 300 SVAAAKRLKLLLPAPVACTSLEYNIRVYCLDTHDALKEVVQLEKQGGQIQPRVLFH 359
QY 703 KDSYHNLRLSLHDVPSSLMKSLVSYQEIPIFYHWNQYRHLCTFTFLRVSPSTSLA 762
Db 360 KDSYHNLRLSLHDVPSSLMKSLVSYQEIPIFYHWNQYRHLCTFTFLRVSPSTSLA 419
QY 763 CKLWVWQVGGQGSFISFNITKDTFPAELLALSEAGVPALVGPSPAKIPFLIRQKIIS 822
Db 420 CKLWVWQVGGQGSFISFNITKDTFPAELLALSEAGVPALVGPSPAKIPFLIRQKIIS 479
QY 823 SLDPCCRRGADWRTIAQKLHLDLSHLSFFASKPSPMTAMILNLWEARHPNGNLSQLAAVA 882
Db 480 SLDPCCRRGADWRTIAQKLHLDLSHLSFFASKPSPMTAMILNLWEARHPNGNLSQLAAVA 539
QY 883 GLGQPDAGLFT-VSEAC 899
Db 540 GTXPAGRWLLSQSCSEAC 557

RESULT 4
US-09-306-902A-6
; Sequence 6, Application US/09306902A
; Patent No. 6277585
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/306,902A
; FILING DATE: 07-May-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341

[illegible]

COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: 2008/08/08,982
APPLICATION NUMBER: US/08/808,982
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC96-217
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 943 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-808-982-7

CORRESPONDENCE ADDRESS:
 ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 STREET: 268 BUSH STREET, SUITE 3200
 CITY: SAN FRANCISCO
 STATE: CALIFORNIA

Db 598 VLSFVTCGPTGLLLCRPVLVTPHCAEVIAGDWIFOLKTAHQGHWE-EVVTLDDETLN 656
Qy 614 HLYYCOLEASACYVTEOLGRFALVGEALSVAARLKLKLLFAPVACTSLEYNIRVYCLH 673
Db 657 TPCYCOLEAKSCHILLDQGLYVFTGESYSRAVKRLQLAIFAPALCTSLSEYLRVYCLE 716
Qy 674 DTHDALKEVVOLKQGLQIEPRVLHFQKDSYHNLRSLSDHVPSSLWKSLLVSYOEIP 733
Db 717 DTPAALKEVLELERTLGGYLVEEPKTLFKDSYHNLRSLSDHDPHAWRSKLLAKYOEIP 776
Qy 734 FYHIWNGTORVHCTFTLERVSPSTSLACKLWVQVGEQGQSINFINITKOTRFAELL 793
Db 777 FYHWNGSQKALHCTFTLERHSLASTFTCKVCVRQVEGEGQIFQLHTTLA-ETPAGSLD 835
Qy 794 ALESEAGVPA--VGSAPKIPFLIRKISLSDPBCRRGADWETLAKKHLDSHLSFFA 851
Db 836 ALCSAPGNAATTOLGPYAFKIPLSIRKICNSLDAPNSRGNDWRLLAQKLSMDRYLYNFA 895
Qy 852 SKPSPTAMILNWEARHPFNGLSQAALAAVAGLQPDAGLFTVSEAC 899
Db 896 TKASPTGVILDWEARQQDDGLNSLASALEMCKSEMVLVAMTTDGDG 943

RESULT 6

US-09-306-902A-7

; Sequence 7, Application US/09306902A

; Patent No. 6277585

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; Leonardo, E. David

; Hink, Lindsey

; Masu, Masayuki

; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/306,902A

; FILING DATE: 07-May-1999

; CLASSIFICATION: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: OSMAN, RICHARD A

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: UC96-217

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 343-4341

; TELEFAX: (415) 343-4342

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 943 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: not relevant

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-306-902A-7

Query Match

Best Local Similarity 52.3%; Pred. No. 2e-227;

Matches 496; Conservative 144; Mismatches 226; Indels 82; Gaps 19;

Qy 9 PALLGIVLAALRWGRGAQSQATVANPVPGANPDLLPHFLVEPDDVYIVKKNPVLVCKAV 68

Db 21 PSLAGI-----DSGAQ-----GLPDSFSPAPAEQLPFLLEPEDAYIVKNPVELHCRAF 70
Qy 69 PATQIFPKONGEWYRQVDHVIERTDSSGSEPTMEVRINVSQRQVEKVFGLYEYWCQVFA 128
Db 71 PATQIFYKNGEWYSQKGVHTVQESLDEATGLRIREVQIEVSQRQVEELFGLYEWQVFA 130
Qy 129 WSSSGITKSKAYIRIARLRKNFROEPLAKEVSLQGIIVLPCRPEGIPPAEVLWRNED 188
Db 131 WSSSGITKSRAYIRIAYLRKNFQDEPLAKEVPLDHEVLLQCRPEGVPVFAEVLWKNED 190
Qy 189 LVDSLDPNVYITREHSLVWRQARLADANTYCVAKNIVAPRRRSASAAVIVYVNGWSTM 248
Db 191 VIDPAQDTNELLTDHNLIIROARLSDTANTYCVAKNIVAKRRSTTACVIVYVNGWSSW 250
Qy 249 TEMSVCSASCGRGWKRSRSTNPAPLNGGAFCEGQNVHRTVSSLLVSDVSGSWPSKWK 308
Db 251 AEWSPCSNRGCGWKRTCTNPAPLNGGAFCEGQ--ACOKTACTTVCVVDGAWTEWSKW 309
Qy 309 SACGLDCTHWSRECSDPAPRNGEECGQTDLTRNCTSDLCV---HSASGPE----- 358
Db 310 SACSTECARWSRECMAPPQNGRCDCGTLSDSKNCTDGLCVLNQRTLNDPKSRPLEPS 369
Qy 359 -DVALYVGL--IAVAVCLVLLLVILVYCRKKEGLSDSDVADSS--ILTSGQPVSIKPSKA 415
Db 370 GDVALYAGLVAVFVVLAVLAVMGVIVYRNRCDFTDITDSSAALTGGFHFVNFKTARP 429
Qy 416 DNPHLL--TIQDLSLTTT--TYQGSCLPRQDGPSPKFQLTNGHLLSPL----- 459
Db 430 SNPQLLHPSAPPDLTASAGIYRGPVVALQDS--ADKIEMTNSPLLDPLSLKIKYVDSSTI 488
Qy 460 --GGG-----RUTLHHSPTSSEAEFVSRLSTQNYPRSRPRG 494
Db 489 GSGAGLADGADLLGLVPPGTYPGDFSRDTHFLHRS-----ASLSGQ--HLLGLPRD 538
Qy 495 TSNNYTGTFNFGRLMIPMTGISLILPPDAIPRGKIYEIVTLHKPEDVRLPLA--GCOT 553
Db 539 PSSSVSTFCGLGRLTIPGTGVSLLVPGNAIFQKGFYDLYLRINKTEST-LPISEGSQT 597
Qy 554 LLSPIVSCGPPGVLLTRPVILAMDHCCEPSPDSMSLRLKQSCGSGWEQDVHLGEEAPS 613
Db 598 VLSFVTCGPTGLLLCRPVLVTPHCAEVIAGDWIFOLKTAHQGHWE-EVVTLDDETLN 656
Qy 614 HLYYCOLEASACYVTEOLGRFALVGEALSVAARLKLKLLFAPVACTSLEYNIRVYCLH 673
Db 657 TPCYCOLEAKSCHILLDQGLYVFTGESYSRAVKRLQLAIFAPALCTSLSEYLRVYCLE 716
Qy 674 DTHDALKEVVOLKQGLQIEPRVLHFQKDSYHNLRSLSDHVPSSLWKSLLVSYOEIP 733
Db 717 DTPAALKEVLELERTLGGYLVEEPKTLFKDSYHNLRSLSDHDPHAWRSKLLAKYOEIP 776
Qy 734 FYHIWNGTORVHCTFTLERVSPSTSLACKLWVQVGEQGQSINFINITKOTRFAELL 793
Db 777 FYHWNGSQKALHCTFTLERHSLASTFTCKVCVRQVEGEGQIFQLHTTLA-ETPAGSLD 835
Qy 794 ALESEAGVPA--VGSAPKIPFLIRKISLSDPBCRRGADWETLAKKHLDSHLSFFA 851
Db 836 ALCSAPGNAATTOLGPYAFKIPLSIRKICNSLDAPNSRGNDWRLLAQKLSMDRYLYNFA 895
Qy 852 SKPSPTAMILNWEARHPFNGLSQAALAAVAGLQPDAGLFTVSEAC 899
Db 896 TKASPTGVILDWEARQQDDGLNSLASALEMCKSEMVLVAMTTDGDG 943

RESULT 7

US-08-808-982-8

; Sequence 8, Application US/08808982

; Patent No. 5932271

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; APPLICANT: Leonardo, E. David

; APPLICANT: Hink, Lindsey

; APPLICANT: Masu, Masayuki

```
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-306-902A-8

Query Match 6.1%; Score 294; DB 2; Length 102;
Best Local Similarity 56.4%; Pred. No. 4.1e-20;
Matches 57; Conservative 16; Mismatches 28; Indels 0; Gaps 0;

Qy 609 EEAPSHLYYQLEASACVVFTEQIGRFALVGEALSVAAAKRLKLLLPAPVACTSLEYNIR 668
Db 2 EETLNTPCYXQLEPRACKILLDQLTGYVFTGESYSRAVKRLQLAVFAPALCTSLYSRLR 61

Qy 669 VYCLHDTHDALKEVVQLEKQGGOLIOEPVRLHFKDSYHNL 709
Db 62 VYCLEDPVALKEVLELRTIGGYLVEEPKPLMPKDSYHNL 102

RESULT 8
US-09-306-902A-8
; Sequence 8, Application US/09306902A
; Patent No. 6277585
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindeay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1172 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1172 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
```

US-08-313-288B-19

Query Match 5.7%; Score 273.5; DB 1; Length 1172;
 Best Local Similarity 28.9%; Pred. No. 2.1e-16;
 Matches 74; Conservative 27; Mismatches 110; Indels 45; Gaps 8;
 QY 209 RQARLADTANYTCAKNTIARERRASAA-VIVYVNGWSTWTSVCSASCGRGWQKESR 267
 DB 403 QRGSRCDVTSNCLGFSIOTACSLSKCDTRIROGGWHSWPSWSSCVTCGVGNITIR 462
 QY 268 SCTNPAPLNGAFCEGQNVHRTVSSLLVSDGWSWPSWKSACGLDCT---HWSRRECS 324
 DB 463 LCNSPVPQMGKNGKNGKSGRETAKCQAGPCPDGRWSPWSPACTVTTCAGGIRERTVON 522
 QY 325 DPARNGEEECQGTDLDRNCTSDLCVHSASGPDVALYVGLIAVAVCLVLLVLLVY 384
 DB 523 SPEFYGGKACGVQERQMCNKRSC-----PVDGCLSNPCFFGAQC----- 564
 QY 385 CRKKEGLSDVADSSILTSFGQPVSI--KPSKADNPHLLTIQDLSLTTT-----TYQ 434
 DB 565 -----SSFPDGS-WSCGFCVGFNGTHCEDLDECALVPDLCSTSKVPCVNTQP 615
 QY 435 GSLC-----PRODGPSP 446
 DB 616 GFHCLPCPPRYRGNQP 631

RESULT 10

US-08-985-526-3
 ; Sequence 3, Application US/08985526
 ; Patent No. 6080728
 ; GENERAL INFORMATION:
 ; APPLICANT: Mixson, James A
 ; TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA
 ; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
 ; TITLE OF INVENTION: THERAPY
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Connolly, Bove, Lodge, & Hutz
 ; STREET: 1220 Market Street, P.O. Box 2207
 ; CITY: Wilmington
 ; STATE: Delaware
 ; COUNTRY: U.S.A.
 ; ZIP: 19899
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/985,526
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/608,845
 ; FILING DATE: 16-JUL-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McMorrow Jr., Robert G
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (302) 658-9141
 ; TELEFAX: (302) 658-5613
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 441 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; US-08-985-526-3

Query Match 5.2%; Score 249; DB 3; Length 441;
 Best Local Similarity 25.9%; Pred. No. 9e-15;
 Matches 87; Conservative 35; Mismatches 114; Indels 100; Gaps 15;
 QY 75 FKCNQEW-----VRQVDHVIERTDGSSEPTM-----EVRINVSQQ---V 113

DB 132 FRQGGWHSWPSWSSCVTCGDGVITRITLNSPSPQMGKPCGEARETKACKKDACPI 191
 QY 114 EKVFLEBYWCQCWVAMSSSGTTKQKAYIRIARLRKNPEQPLAKEVSLQGIIVLPCRPP 173
 DB 192 NGWGPWSPWDCSVTCGGGVQKSRCLV---DSRMTEENKELANELR-----RPP 239
 QY 174 ----EGIPPAEVEWLNRD-LVDPSPDPNVVITREHSLVVRQARLADTANYTCAKNTI 228
 DB 240 LCYHNG-----VOYRNNEEWTVDSCTE-----CHQONSVT 269
 QY 229 RRSASAAVIVVYNG-----GWSTWTSVCSASCGRGWQKESRSC 269
 DB 270 ICKKVCPIIMPCSNATVPDGECCPRCWPSDSADDDGWSWSEWTSCTSCGNGIQORGRSC 329
 QY 270 TNPAFLNGAFCEGQNVHRTV---SSLLVSDGWSWPSWKSACGLDCT---THWSRE 322
 DB 330 DS---LNNR---CEGSSVQTRTCHIQCCKRFPQDGGHSHWSPWSSCVTCGDGVITRITL 384
 QY 323 CSDPAPRNGEEECQGTDLDRNCTSDLC-VHSASGP 357
 DB 385 CNSPQMGKPCGEARETKACKKDACPIINGWGP 420

RESULT 11

PCT-US93-01652-1
 ; Sequence 1, Application PC/TUS9301652
 ; GENERAL INFORMATION:
 ; APPLICANT: Bouck, Noel P.
 ; APPLICANT: Polverini, Peter J.
 ; APPLICANT: Good, Deborah J.
 ; APPLICANT: Frazier, William A.
 ; TITLE OF INVENTION: Method and Composition for
 ; TITLE OF INVENTION: Inhibiting Angiogenesis
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Filton, Fallon, Lungmus & Chestnut
 ; STREET: 100 South Wacker Drive, Suite 960
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60606-4002
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/01652
 ; FILING DATE: 19930222
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/841,656
 ; FILING DATE: 24-FEB-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/464,369
 ; FILING DATE: 12-JAN-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fentress, Susan B.
 ; REGISTRATION NUMBER: 31,327
 ; REFERENCE/DOCKET NUMBER: 92005-PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (312)-456-8000
 ; TELEFAX: (312)-456-7776
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 239 amino acids
 ; TYPE: AMINO ACID
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: peptide
 ; PCT-US93-01652-1


```

Query Match      5.1%; Score 246; DB 5; Length 239;
Best Local Similarity 31.0%; Pred. No. 6.2e-15;
Matches 48; Conservative 23; Mismatches 66; Indels 18; Gaps 3;

QY 207 VVQRLADTANYTCVAKNIIVAR-----RRSASAAVIVVNGGWTWTWWSVCASASC 258
Db 88 IQQGRGSCDSLNNRCEGSSVQTRTCHTQECDKRFKQ-----DGGSHWSPWSSCSVTC 140
QY 259 GRGQKRSRCTNPAPLNGGAFCEGQNVHDTVSSLLVSDGWSNPKWSKWSACGLDC--- 315
Db 141 GDGVITRILCNPSPOQNGPLPCEGEARETKACKKDACPINGGWPSPMDICSVTCGGG 200
QY 316 THWRSRECSDPAPRNGGECOGTDLDRNCTSDLC 350
Db 201 VQKRSRLCNPAPOFGGLDCVGDVTENQICNKKQDC 235

RESULT 12
US-08-313-288B-20
; Sequence 20, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-20

Query Match      5.1%; Score 246; DB 1; Length 1170;
Best Local Similarity 30.3%; Pred. No. 8.8e-14;
Matches 47; Conservative 24; Mismatches 66; Indels 18; Gaps 3;

QY 207 VVQRLADTANYTCVAKNIIVAR-----RRSASAAVIVVNGGWTWTWWSVCASASC 258
Db 399 IQQGRGSCDSLNNRCEGSSVQTRTCHTQECDKRFKQ-----DGGSHWSPWSSCSVTC 451
QY 259 GRGQKRSRCTNPAPLNGGAFCEGQNVHDTVSSLLVSDGWSNPKWSKWSACGLDC--- 315
Db 452 GDGVITRILCNPSPOQNGPLPCEGEARETKACKKDACPINGGWPSPMDICSVTCGGG 511
QY 316 THWRSRECSDPAPRNGGECOGTDLDRNCTSDLC 350

```

```
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/065,544
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: 60/081,057
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1395
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-540-245A-15

Query Match      4.9%; Score 234; DB 3; Length 1395;
Best Local Similarity 20.5%; Pred. No. 1.6e-12;
Matches 188; Conservative 103; Mismatches 262; Indels 364; Gaps 46;

Qy 4 RPGLWPALLGLVLAARLGRSGAQSATVANVPGA-----NPDLLPHFLVPEPDIYVKN 58
Db 28 RMWLLPAMLLVLA-----SNGLPVRGQYQSPRIEH-----PTDLVVKKN 70
Qy 59 KPVLLVCK--AVPATQIFKCKNGEVRQVDHVIERSIDGSGEPTMEVRINVR-----110
Db 71 EPATLNCKVEKPEPTI-----EWFK-----DGEF-----VSTNEKSKHRVQF 108
Qy 111 -----QQVEKVFLEBYWCQVAMSSSGTTKQKAYIRIARLKNFEQEPPLAKEV 160
Db 109 KDGALFFVRTWQKKEQDGGYW--CVAKNRVGOAVSRHASLQIAVLDRDFRVEPKTRV 166
Qy 161 SLEOGIVLPCPPPGIGIPAEVWURN-----EDLVDSL--DPNYYITREHSLVVRQARLA 214
Db 167 AKGETALCEGPKPGIPEPTLIMIKDGVPLDLDKAMSGFASRRVRIVDGNNLLISNVEPI 226
Qy 215 DTANYTCVAKNIVARRRSASAAVLYVN-----GWSWTWTEWSVCSACGR 260
Db 227 DEGNKICIAQLNVGTRESSYAKGLIVQKPFMKPEKPDQVMLYGQTATF-----HCSV 278
Qy 261 G-----WQK-----RSRSCNTPAPLNGGAF--CEGQN-----285
Db 279 GGDPPPKVLWKKEBGNIPVSRARILHDEKLSLEISNITPTDEGTVVCAHNNVGOISARAS 338
Qy 286 --VH-----DRTVSSLLVSDG-----SWSP-----WSKWSACGL--DCTHWSRE 322
Db 339 LIVHAPNFTKRSNKKVGLGVVQLPCMASGNPPSVFMTKEGVSTLMPFNSSHGQYV 398
Qy 323 CSDFAPRNGEECOGTDLTDRNCTSDLVCHSASGPEVALYVGLIYAVAVCLVLLVLL 382
Db 399 AADGT-----LQITDV-----RQDEGYI-----VCSAF-----422
Qy 383 VYCRKKEGLSDVDSSILTSGFQPVSIKPSKADN--PHLLTIQPDLSITTTTQSGSL---437
Db 423 -----SVVDSSTVR-----VFLQVSSVDERPPIIIQGP--ANQTLPGKSVATL 464
Qy 438 -CPRQDGPSPKFQ-LTNGHLLSLPGLGGRHTLHSSPTSEAEFVSRLSTQNYFRSLPRGT 495
Db 465 PCRAATGNPSRIKFWFDGHAVQ--AGNRYGIIQGSLSLRVDDLQLSDSGTCTASGERGE 522
Qy 496 SNMTYGTGFLGRLMIPNTGISILLIPDAIPRGIYIYVTLHKPEDEVPLPLAGCOTLL 555
Db 523 TS-----WAATLTVEKPGSTSLHRAA-----543
Qy 556 SPITVSGPPGVLLTRPVILAMDHCGEPPSPDSWSLR-LKQSCSGS-----599
Db 544 DPSTYPAPPGT---PKVLNV-----SRTSISLRWAKSKEKPGAVGPIIYTVVEFSPD 593
Qy 600 -----NEQDVLHGLEAPSHUYCOLEASACYVF-----TEQLGR 634
Db 594 LQTGWIVAAHRVGD---TQVTISGLTPGTSYFVLVRAENTQGISVPSGLSNVIKTIADF 650
Qy 635 FALVGEALSVA---AAKRLKLLLPAPVACTSLEYNIRVYVCLHDTHDALKEVVOLEKQLG 690
Db 651 DAASANDLSAARTLLTGKSVELIDASAINASAVRLE---WMLHVSAD-----BKYVE 699
```

```
Qy 691 GQLIQEPRLVLFKDS-----YHNLRL-----SIHNV 716
Db 700 GLRI-----HYKDSVPSAQYHSITVMDASAEFVWGNLKKYTKYEFPLTPFFETIEGQ 753
Qy 717 PSSLWKSXLLVSYOEIP 733
Db 754 PSN---SKTALTYEDVP 767

RESULT 15
US-09-540-245A-18
; Sequence 18, Application US/09540245A
; Patent No. 6270984
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey
; APPLICANT: Kid, Thomas
; APPLICANT: Brosse, Katja
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions
; FILE REFERENCE: 898-031-3
; CURRENT APPLICATION NUMBER: US/09/540,245A
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/065,544
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: 60/081,057
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 18
; LENGTH: 1651
; TYPE: PRT
; ORGANISM: human
US-09-540-245A-18

Query Match      4.5%; Score 216.5; DB 3; Length 1651;
Best Local Similarity 20.7%; Pred. No. 1e-10;
Matches 213; Conservative 139; Mismatches 381; Indels 295; Gaps 54;

Qy 41 DLLPHFLVPEPDIYVKNKPVLLVCKAV--PATQIFPKCNGEWWVRQVDHVIERSIDGSSG 98
Db 65 DFPFRIVEHPSDLIVSGEPATLNCKAEGRPPTI-----EWYKGERVETDKDDPRSH 118
Qy 99 EPTME-----VRINVSROQVEKVFLEBYWCQVAMSSSGTTKQKAYIRIARLKNFE 152
Db 119 RMLPSGSLFFLLIVHGRKSRP-----DEGVYCVARNYLGEAVSHNASLEVALRDDR 173
Qy 153 QEPLAKEVLSLEQIVLPCRPPEGIPPAEVEWLRNEDLVDPSPDPNVYIT-REHSLVVRQA 211
Db 174 QNPSDVNVAVGEPAVMECQPPRGHPPTISWKKDGSPLD---DKDERITIRGGKLMITYT 230
Qy 212 RLADTANYTCVAKNIVARRRSASAAVLYVNGW-----STWTEWSV--CSASCGR 261
Db 231 RKSDAGKYVCGVTNMGVERSEVAELTVLERPSFVKRPSNLAVTVDDSDAEFKCEA---RG 287
Qy 262 -----WOK-----RSRSCNTPAPLNGGAFCEGQNVHDTVSSLLV 296
Db 288 DPVPTVRNRKDGELPKSRYEIRDDHTLKRKYT--AGDMGSYTCVAENMVGAESAATL 345
Qy 297 SVDGS-----WSPMSKWSACGLDCTHWRSECSDPAP-----RNGGEEC-----OQTDL 340
Db 346 TVQEPHFVVKPDDQVVALGRVTY--FQCEATGNPQPAIFWRREGSQNLLFSYOPPOSSSR 404
Qy 341 DTRNCTSDLVCHSASGPEVALYVGLIYAVAVCLVLLVLLVLYCRKKEGLSDVDADSSI 400
Db 405 FSVSQTGDLTIITNVQ-RSDVGYY-----ICQTL-----NVA-GSI 437
Qy 401 LTSFGFQPVSIKPSKADNPHLLTIQPDLSITTTTQSGSL---CPRQDGPSPKF-----448
Db 438 ITKAY--LEBTVDIARPPPIRQGVNQTVAVDGTFLVSCVATGSPFTILWRKDGVLV 495
Qy 449 -----QLTNGHL---LSPILGG--GRHTLHSSPTSEAEFVSRLSTQNYFRSLPRGTS 496
Db 496 STODSRIKQLENGVLOIRYAKLGDGTGYTCIASTPSGEA-----534
```

Thu Oct 7 10:58:55 2004

497	Qy	QY	NWYGTGTFNLGRLMWNPTGHSILIP	----	PDATPRCKIYEVILTHKPE	--DV--	----	RLP	544
535	Db	DB	--TWSAY	-----	IEVQBEFGVOPPRPTFNLPISAP	-----	SKPEVTDVSRNVT	578	
548	Qy	QY	LAGCOTLLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRKTKQSCGSGWEQDVHLH	607					
579	Db	DB	LSWQPNLNS	-----	CATPTSIIIEAFSHA	----	SGSSMQTVAENVKTETSAIKGL	624	
608	Qy	QY	GEEARSHLYYCOLEASACYVETE	----	QLGRFALVGEALSVA	----	KRLK	655	
625	Db	DB	---KPNATYFLVRAANAAGHSDPSQISDPVKYQDVLPFGQVDHQRQVORELGNVHLH	681					
656	Qy	QY	APVACTSLEYNIRVYCLHDTHDALKEVVQLEKQJGGQLIQEPRVLRFKQSYHNLRLSIHD	715					
682	Db	DB	NFTVLSSSIEV	-----	HWTVDQSQYIQ	-----	GKILYRPSGANHGES	727	
716	Qy	QY	VPSSLWKSLL	-----	VSYQ	----	ELPFVHIWNGTQRYLHCTFLIERVSPSTSLACKLMV	767	
728	Db	DB	VRTPAKNSWIPDLRCKGNYIKARPFFNEPQGDSEIKFAKTL	784					
768	Qy	QY	MOVEGDGQGSFINFN	----	ITKDRTRFAEL	-----	LALSEAGVPALVGPSAFK	817	
785	Db	DB	SKNDGNGTALLVSMQPPEDTQNGMVQKYKVMCLGNTRYHINKTVDGSSFFSVVIFELV	843					
818	Qy	QY	QXIIISLDPPCRGA	----	DWRTILAKLHDSH	-----	LSPTASKPSTA	858	
844	Db	DB	PGIRYSVEVAASTGAGSGVKSEPOFIQLDAGHNVPVSPEDQVSLAQQISDVVKVPATFAGI	903					
859	Qy	QY	-----MILMLWEARHFP	----	NGNLSQLAA	-----	AVAGLGQPD	888	
904	Db	DB	GRACWIIILMVSIWLYRRKKRNGLTSTYAGIRKVPSTFTPTPTVYORGGEAVSSGRP	962					
889	Qy	QY	AGLTVSE	896					
963	Db	DB	-GLLNISE	969					

Search completed: October 6, 2004, 18:20:40
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 12:45:02 ; Search time 88 Seconds
(without alignments)
2886.480 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVPGLWPLLGLVLAWL.....AVAGLQPDAGLFTVSAEC 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4787	100.0	899	5	AAU79939 Human UNC
2	4698.5	98.2	898	5	AAU85403 Human pro
3	4688.5	97.9	898	5	AAU97899 Human net
4	4545.5	95.0	898	2	AAW78898 Rat UNC-5
5	4545.5	95.0	898	5	AAU10543
6	4545.5	95.0	898	5	AAU97900
7	4434	92.6	943	4	AAU79128 Human pro
8	4379.5	91.5	842	5	AAU74818 Human REP
9	2789	58.3	556	2	AAW78899 Human UNC
10	2673.5	55.8	931	4	AAU50691 Human UNC
11	2673.5	55.8	931	7	ADEG3098 Human pro
12	2673.5	55.8	932	4	ABG11551 Novel hum
13	2497	52.2	945	7	ADEG3096 Rat Prote
14	2490.5	52.0	946	5	ABG61795 Novel UNC
15	2490.5	52.0	943	2	AAW78900
16	2484	51.9	933	5	AAO18734 Human NOV
17	2484	51.9	933	5	AAO18735 Human NOV
18	2479	51.8	945	4	AAU12244 Human PRO
19	2479	51.8	945	6	ABO17688 Novel hum
20	2479	51.8	945	6	ABU80942 Human PRO
21	2479	51.8	945	6	ABU6642 Human PRO
22	2479	51.8	945	6	ABU59723 Novel sec
23	2479	51.8	945	6	ABO24913 Human sec
24	2479	51.8	945	6	ABU66918 Human sec
25	2479	51.8	945	6	ADA45665 Novel hum

26	2479	51.8	945	6	ADA76096 Human PRO
27	2479	51.8	945	6	ADA18746 Human PRO
28	2479	51.8	945	6	ADA61369 Homo sapi
29	2479	51.8	945	6	ADB19154 Novel hum
30	2479	51.8	945	6	ADB27695 Human PRO
31	2479	51.8	945	6	ADA86174 Novel hum
32	2479	51.8	945	6	ADB15738 Human PRO
33	2479	51.8	945	6	ADA47524 Human PRO
34	2479	51.8	945	6	ADA67319 Human PRO
35	2479	51.8	945	6	ADB30326 Human PRO
36	2479	51.8	945	6	ADA85622 Novel hum
37	2479	51.8	945	6	ADA96834 Human PRO
38	2479	51.8	945	6	ADA79138 Human PRO
39	2479	51.8	945	6	ADA87277 Novel hum
40	2479	51.8	945	6	ADB16479 Human PRO
41	2479	51.8	945	6	ADA91571 Novel hum
42	2479	51.8	945	6	ADB14634 Human PRO
43	2479	51.8	945	6	ADB18595 Novel hum
44	2479	51.8	945	6	ADA93810 Human PRO
45	2479	51.8	945	6	ADB19706 Novel hum

ALIGNMENTS

RESULT 1
AAU79939
ID AAU79939 standard; protein; 899 AA.
XX AC
XX AAU79939;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human UNC5-like protein NOV1.
XX
KW Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;
KW cell signal processing; metabolic pathway modulation; cancerous tissue;
KW antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;
KW chromosome 13.
XX
OS Homo sapiens.
XX
PN WO200229038-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031377.
XX
PR 04-OCT-2000; 2000US-0237862P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Herrmann JL, Rastelli L, Shinkets RA;
XX WPI; 2002-340104/37.
XX N-PSDB; ABK49422.
XX
PT Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for
treating cardiomyopathy, atherosclerosis, and cancer.
XX
PS Claim 1; Page 9; 180pp; English.

CC The present invention relates to a new NOVX polypeptide having a 900
CC (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or 331 (NOV6)
CC residue amino acid sequence, as given in the specification. The novel
CC polypeptide, and its encoding polynucleotide, are used to treat
CC cardiomyopathy, atherosclerosis, cancer or a disease related to cell
CC signal processing and metabolic pathway modulation, in a human. Detecting
CC the polypeptide or polynucleotide is useful for identifying cancerous
CC tissue. The antibody can be used to treat diabetes or cancer. The host
CC cells can be used to produce non-human transgenic animals useful in drug
CC screening. The present amino acid sequence is that of the human UNC5-like
CC protein NOV1 of the invention. This sequence is encoded by the human UNC5

CC such as Huntington's disease and Gilles de la Tourette's syndrome. The
CC nucleic acid is useful in gene therapy. The present sequence represents a
CC NOVX protein
XX
SQ Sequence 898 AA;

Query Match 98.2%; Score 4698.5; DB 5; Length 898;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 888; Conservative 2; Mismatches 7; Indels 3; Gaps 3;

Qy 1 MAVRPLGALLGIVLAWLKSGAQQAQATVANPVGANPDLLPHFLVPEDEVYVKNKP 60
Db 1 MAVRPLGALLGIVLAWLKSGAQQAQATVANPVGANPDLLPHFLVPEDEVYVKNKP 60
Qy 61 VLLVCKAVPATQIFPKNGEWRQVDHVIERTSGSSCEPTMEVRIINVSROQVEKVGLE 120
Db 61 VLLVCKAVPATQIFPKNGEWRQVDHVIERTSGSSGLPTMEVRIINVSROQVEKVGLE 120
Qy 121 EYWCQCVAMSSSGTTKQKAYIRIARLKNFEQEPPLAKEVLEQIVLPCRPPEGIPPAE 180
Db 121 EYWCQCVAMSSSGTTKQKAYIRIARLKNFEQEPPLAKEVLEQIVLPCRPPEGIPPAE 180
Qy 181 VEWLRNEDLVDPSPDNPVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
Db 181 VEWLRNEDLVDPSPDNPVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
Qy 241 VNGGWSWTWVSCVSCGRCQKRSCTNPAPLNGGAFCEGQVNDRTVSSLLVSDG 300
Db 241 VNGGWSWTWVSCVSCGRCQKRSCTNPAPLNGGAFCEGQVNDRTVSSLLVSDG 300
Qy 301 SWSPWKSACGLDCTHWRSECDPAPRNGEGECQGTDLDTNRCTSDLCVHSASGPEV 360
Db 301 SWSPWKSACGLDCTHWRSECDPAPRNGEGECQGTDLDTNRCTSDLCVHSASGPEV 360
Qy 361 ALYVGLTAVAVCLVLLVLLVLYCRKEGLSDVADSSILTSQFQVSIKPSKADNPHL 420
Db 361 ALYVGLTAVAVCLVLLVLLVLYCRKEGLSDVADSSILTSQFQVSIKPSKADNPHL 420
Qy 421 LTIQPDLS-TTTTYQGLSCPRQDQSPKPOLTNGHLLSPGCGGHTLHSSPTSEAEFFV 479
Db 421 LTIQPDLS-TTTTYQGLSCPRQDQSPKPOLTNGHLLSPGCGGHTLHSSPTSEAEFFV 479
Qy 480 SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRIMIPNTGISLLIPDAPRGIKVIYITLH 539
Db 480 SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRIMIPNTGISLLIPDAPRGIKVIYITLH 539
Qy 540 KPEDVRLPIAGCQTLSPVSCGPPGVLLTRPVILAMDHGCEPSPDSWSRLKKQCEGS 599
Db 540 KPEDVRLPIAGCQTLSPVSCGPPGVLLTRPVILAMDHGCEPSPDSWSRLKKQCEGS 599
Qy 600 WEQDVHLGGEAPSHLYCYOLEASACVFTQELGRFALVGEALSVAAAKELKLLFAPVA 659
Db 600 WE-DVHLGGEAPSHLYCYOLEASACVFTQELGRFALVGEALSVAAAKELKLLFAPVA 658
Qy 660 CTSLEYNIRVYCLDTHDALKEVQLEKQGGQILQEPVRLHFKDSYHNLRLSHDVPSS 719
Db 659 CTSLEYNIRVYCLDTHDALKEVQLEKQGGQILQEPVRLHFKDSYHNLRLSHDVPSS 718
Qy 720 LMKSKLLVSQEIPIFFHINNGTORYLHCTTLTBRVSPSTSDACKLWVQVGGQGSFSI 779
Db 719 LMKSKLLVSQEIPIFFHINNGTORYLHCTTLTBRVSPSTSDACKLWVQVGGQGSFSI 778
Qy 780 NFNIKTDTFAELLALASEAGVAPVAFKIPFLTRQKIISSLPCCRGADWTLAQ 839
Db 779 NFNIKTDTFAELLALASEAGVAPVAFKIPFLTRQKIISSLPCCRGADWTLAQ 838
Qy 840 KLHLDLSLFFASKPSPTAMILNLWEARHPFNGNLSQAAAAGLQOPDAGLFTVSEAC 899
Db 839 KLHLDLSLFFASKPSPTAMILNLWEARHPFNGNLSQAAAAGLQOPDAGLFTVSEAC 898

RESULT 3
AAU97899

ID
XX
AC
XX
XX
DT
XX
XX
DE
XX
KW
KW
KW
XX
OS
XX
XX
FH
FT
FT
FT
FT
FT
FT
FT
FT
FT
XX
PN
XX
XX
PD
XX
PF
XX
PR
XX
PA
XX
XX
PI
XX
DR
XX
DR
XX
PT
PT
PT
PT
XX
PS
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
SQ

AAU97899 standard; protein; 898 AA.
AAU97899;
27-AUG-2002 (first entry)
Human netrin binding membrane receptor UNC5H-1 protein.
Netrin binding membrane receptor; receptor; UNC5H-1; human; neurotropic;
neuroprotective; cytosolic; antiparkinsonian; cerebroprotective; cancer;
central nervous system; CNS; stroke; Parkinson's disease;
multiple sclerosis; Alzheimer's disease.
Homo sapiens.
Key Location/Qualifiers
Domain 152..223
/note= "Immunoglobulin domain "
Domain 247..294
/note= "Thrombospondine type 1 domain "
Domain 302..348
/note= "Thrombospondine type 1 domain"
Region 361..382
/note= "Transmembrane region"
Domain 495..598
/note= "ZU5 domain"
Domain 817..897
/note= "Death domain"
WO200233080-A2.
25-APR-2002.
15-OCT-2001; 2001WO-EF011891.
16-OCT-2000; 2000US-0240061P.
(FARB) BAYER AG.
Koehler RH;
WPI: 2002-463314/49.
N-PSDB; ABK32891.
Novel human netrin binding membrane receptor polypeptide and
polynucleotides for identifying modulating agents useful in treating
diseases e.g. Parkinson's disease, multiple sclerosis, stroke,
Alzheimer's disease.
Claim 1; Fig 2; 94pp; English.
This invention relates to the DNA and protein sequences of a novel
purified human netrin binding membrane receptor, UNC5H-1. The DNA
sequence of the invention is useful as a probe for detecting a nucleic
acid encoding the UNC5H-1 protein in a biological sample. The sequences
of the invention are useful to screen for agents which decrease the
activity of the UNC5H-1 protein. The sequences are also useful for
screening agents which regulate (modulate) the activity of the protein of
the invention. A pharmaceutical composition containing the protein of the
invention or a reagent that modulates the activity of the UNC5H-1 protein
may be useful for treating a UNC5H-1 dysfunction related disease such as
cancer or a central nervous system (CNS) disorders (e.g. Parkinson's
disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion
proteins comprising the UNC5H-1 protein are useful for generating
antibodies and for in various assay systems, and the protein can be used
as a bait protein in a two-hybrid assay or three-hybrid assay. The method
of the invention is useful for detecting a coding sequence for the UNC5H-
1 protein. The present sequence represents the human netrin binding
membrane receptor UNC5H-1 protein of the invention
Sequence 898 AA;

Query Match 97.9%; Score 4688.5; DB 5; Length 898;

Best Local Similarity 98.4%; Pred. No. 0;		Matches 886; Conservative 3; Mismatches 8; Indels 3; Gaps 3;	
Qy	1	MAVRPGLWPALLGIVLAANLRGSAQOQSAIVANVPGANPDLLPHFLVEPEDEVYIVKNKP	60
Db	1	MAVRPGLWPALLGIVLAANLRGSAQOQSAIVANVPGANPDLLPHFLVEPEDEVYIVKNKP	60
Qy	61	VLLVCKAVPATQIFPKCNQGEWVRQVDHVIERTSDGSGEPTMEVRINVSQQQKVFGL	120
Db	61	VLLVCKAVPATQIFPKCNQGEWVRQVDHVIERTSDGSGEPTMEVRINVSQQQKVFGL	120
Qy	121	EYWCQCVAWSSGTTKSKAYIRIARLRKNFEQPLAKEVSLQGIPLCPPEGIPPAE	180
Db	121	EYWCQCVAWSSGTTKSKAYIRIARLRKNFEQPLAKEVSLQGIPLCPPEGIPPAE	180
Qy	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIIVARRSAAVIV	240
Db	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIIVARRSAAVIV	240
Qy	241	VNGGWTWTSVCSACGRGWQRKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG	300
Db	241	VNGGWTWTSVCSACGRGWQRKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG	299
Qy	301	SWSPWKSACGLDCTHWRSECDPAPRNGEBCQGTDLTRNCTSDLCVHSASGPEV	360
Db	300	SWSPWKSACGLDCTHWRSECDPAPRNGEBCQGTDLTRNCTSDLCVHSASGPEV	359
Qy	361	ALYVGLIAVAVCLVLLVLLVYCRKKEGSDSVADSSILTSFGFQPVSIKPSKADNPHL	420
Db	360	ALYVGLIAVAVCLVLLVLLVYCRKKEGSDSVADSSILTSFGFQPVSIKPSKADNPHL	419
Qy	421	LTIQPDLS-TTTTQGSICPRQDGPSPKQFQTLNTHLLSLPLGGGRTLHSSPTSEAEFV	479
Db	420	LTIQPDLS-TTTTQGSICPRQDGPSPKQFQTLNTHLLSLPLGGGRTLHSSPTSEAEFV	479
Qy	480	SRLSTQNYFRSLPRGTSNMTYGTNFGRLMIPNTGSLIPDIPDAIPRGKIYEIYTLH	539
Db	480	SRLSTQNYFRSLPRGTSNMTYGTNFGRLMIPNTGSLIPDIPDAIPRGKIYEIYTLH	539
Qy	540	KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPSPDSLSLKQSCGSG	599
Db	540	KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPSPDSLSLKQSCGSG	599
Qy	600	WEQDVILHGEAPSHLYCOLEASACYFTQLGRFALVGEALSVAARLKLILFAPVA	659
Db	600	WE-DVLHGEAPSHLYCOLEASACYFTQLGRFALVGEALSVAARLKLILFAPVA	658
Qy	660	CTSLYNIIRVYCLDTHDALKEVQLEKQGLQIQRVHLHFQDSYHNLRSLIHDVPS	719
Db	659	CTSLYNIIRVYCLDTHDALKEVQLEKQGLQIQRVHLHFQDSYHNLRSLIHDVPS	718
Qy	720	LWKSLLVSYQEIPIFYHWNQTRQYLHCTFLERVSPTSIDLACKLWVQVEGDCQSFSI	779
Db	719	LWKSLLVSYQEIPIFYHWNQTRQYLHCTFLERVSPTSIDLACKLWVQVEGDCQSFSI	778
Qy	780	NFNITKTRFRELIALESEAGVPALVGPSAKIPIELIRQKLISSLDPPCRGADWRTLAQ	839
Db	779	NFNITKTRFRELIALESEAGVPALVGPSAKIPIELIRQKLISSLDPPCRGADWRTLAQ	838
Qy	840	KLHLDLSLSPFASKPSPTAMILNWEARHPNGLSQAALAAVAGLQPDAGLFTVSEAE	899
Db	839	KLHLDLSLSPFASKPSPTAMILNWEARHPNGLSQAALAAVAGLQPDAGLFTVSEAE	898
RESULT 4			
ID	AAW78898		
XX	AAW78898 standard; protein; 898 AA.		
AC	AAW78898;		
XX	25-MAR-2003 (revised)		
DT	21-DEC-1998 (first entry)		
XX			

DE	Rat UNC-5 homologue UNC5H-1.
XX	UNC-5; UNC5H-1; rat; netrin receptor; cell migration; axon guidance;
KW	diagnosis; therapy.
XX	Rattus sp.
XX	Key Location/Qualifiers
FT	Peptide 580..594
FT	/note="peptide used to raise rabbit polyclonal antisera"
XX	WO9837085-A1.
XX	27-AUG-1998.
XX	19-FEB-1998; 98WO-US003143.
XX	19-FEB-1997; 97US-00808982.
XX	(REGC) UNIV CALIFORNIA.
XX	Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasu K;
XX	WPI; 1998-495364/42.
XX	N-PSDB; AAV52940.
XX	Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and
XX	the biopharmaceutical industry.
XX	Claim 1; Page 19-22; 32pp; English.
XX	UNC5H-1 and UNC5H-2 (see AAW78900) are rat homologues of Caenorhabditis
XX	elegans UNC-5 protein. Their amino acid sequences were deduced from
XX	isolated unc5h cDNA clones (see AAV52940 and AAV52942) isolated from an
XX	E18 brain cDNA library. The predicted proteins show similarity with UNC-
XX	5, possess 2 predicted Ig-like domains and 2 predicted thrombospondin
XX	type-1 repeats, a predicted membrane spanning region, and a large
XX	intracellular domain. They are predicted to be involved in cell migration
XX	and axon guidance, and are characterised as receptor proteins for
XX	netrins. Human UNC5H-1 (see AAW78899) and UNC5H-2 (see AAW78901) proteins
XX	are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly
XX	from transfected host cells. The invention also provides unc-5
XX	hybridisation probes and primers, vertebrate UNC-5-specific binding
XX	agents such as specific antibodies, and methods of making and using the
XX	subject compositions in diagnosis (e.g. genetic hybridisation screens for
XX	vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate
XX	vertebrate unc-5 gene expression) and in the biopharmaceutical industry
XX	(e.g. as immunogens, reagents for modulating cell guidance, reagents for
XX	screening chemical libraries for lead pharmacological agents, etc.).
XX	(Updated on 25-MAR-2003 to correct PI field.)
XX	Sequence 898 AA;
Qy	Query Match 95.0%; Score 4545.5; DB 2; Length 898;
Db	Best Local Similarity 94.7%; Pred. NO. 0;
Qy	Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;
Db	1 MAVRPGLWPALLGIVLAANLRGSAQOQSAIVANVPGANPDLLPHFLVEPEDEVYIVKNKP 60
Qy	1 MAVRPGLWPALLGIVLAANLRGSAQOQSAIVANVPGANPDLLPHFLVEPEDEVYIVKNKP 60
Db	61 VLLVCKAVPATQIFPKCNQGEWVRQVDHVIERTSDGSGEPTMEVRINVSQQQKVFGL 120
Qy	61 VLLVCKAVPATQIFPKCNQGEWVRQVDHVIERTSDGSGEPTMEVRINVSQQQKVFGL 120
Db	121 EYWCQCVAWSSGTTKSKAYIRIARLRKNFEQPLAKEVSLQGIPLCPPEGIPPAE 180
Qy	121 EYWCQCVAWSSGTTKSKAYIRIARLRKNFEQPLAKEVSLQGIPLCPPEGIPPAE 180
Db	181 VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIIVARRSAAVIV 240
Qy	181 VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIIVARRSAAVIV 240

QY 241 VNGWSTWTEWVSCASCGRGWKRSRSCNTPAPLNGAFCEGQNVHDTVSSLLVSDG 300
 Db |||||
 QY 241 VNGWSTWTEWVSCASCGRGWKRSRSCNTPAPLNGAFCEGQNV-QKTACATLCPVDG 299
 Db |||||
 QY 301 SWSPKWSACGLDCTHWRSECDPAPRNGEBCQGTDLDRNCTSDLCVHSASGPEV 360
 Db |||||
 QY 300 SWSPKWSACGLDCTHWRSECDPAPRNGEBCGADLDRNCTSDLCVHSASGPEV 359
 Db |||||
 QY 361 ALYVGLIAVAVCLVLLVLLVYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL 420
 Db |||||
 QY 360 ALYIGLVAVAVCLFLLLLALGLIYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL 419
 Db |||||
 QY 421 LTIQPDLS-TTTTYQGSCLPRQDGPSPKQFQLTNGHLLSPLGGGRHTLHSSPTSEAEFV 479
 Db |||||
 QY 420 LTIQPDLS-TTTTYQGSCLPRQDGPSPKQFQLTNGHLLSPLGGGRHTLHSSPTSEAEFV 479
 Db |||||
 QY 480 SRLSTONYFRSLPRGTSNNYGTFTNFGRLMIPNTGISLLIPDAIPRGKIYEIYTLH 539
 Db |||||
 QY 480 SRLSTONYFRSLPRGTSNNYGTFTNFGRLMIPNTGISLLIPDAIPRGKIYEIYTLH 539
 Db |||||
 QY 540 KPDEVRLPLAGCQTLISPIVSCGPGVLLTRPVILAMDHCGEPSPDSWSLRKKQSCGS 599
 Db |||||
 QY 540 KPDEVRLPLAGCQTLISPIVSCGPGVLLTRPVILAMDHCGEPSPDSWSLRKKQSCGS 599
 Db |||||
 QY 600 WEQVHLHGEAPSHLYYCQLEASACYYFTBQLGRFALVGEALSVAARLKLILFAPVA 659
 Db |||||
 QY 600 WE-DVLHGLGEPSHLYYCQLEAGACYYFTBQLGRFALVGEALSVAATKRLRLILFAPVA 658
 Db |||||
 QY 660 CTSLEYNIRVYCLDTHDALKEVQLEKQGLIQEPRVLFHFKDSYHNLRSLSDHVPSS 719
 Db |||||
 QY 659 CTSLEYNIRVYCLDTHDALKEVQLEKQGLIQEPRVLFHFKDSYHNLRSLSDHVPSS 718
 Db |||||
 QY 720 LKWSKLLVSYQEIPIFYHWNQYRLHCTFTLERVSPSTSLACKLWVQVGGQGSFSI 779
 Db |||||
 QY 719 LKWSKLLVSYQEIPIFYHWNQYRLHCTFTLERINASTSLACKVWVQVGGQGSFNI 778
 Db |||||
 QY 780 NFNITKDTFAELIALSEAGVPALVGSAPKIFPLIRQKIISLDPCCRGADWRTLAQ 839
 Db |||||
 QY 779 NFNITKDTFAELIALSEAGVPALVGSAPKIFPLIRQKIISLDPCCRGADWRTLAQ 838
 Db |||||
 QY 840 KLHLDLSLFFASKPSATMILNWEARHPNGNLQSLAAVAGLQPDAGLFTVSEAE 899
 Db |||||
 QY 839 KLHLDLSLFFASKPSATMILNWEARHPNGNLQSLAAVAGLQPDAGLFTVSEAE 898
 Db |||||

RESULT 5
 AAU10543
 ID AAU10543 standard; protein; 898 AA.
 AC AAU10543;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Rat netrin receptor UNC5H1 (YSG7) polypeptide.
 XX
 KW YSG7; YSG7; schizophrenia; chronic animal model; LCGU; netrin receptor;
 KW local cerebral glucose utilisation; phosphodiesterase 1-alpha; UNC5H1;
 KW calcium-independent alpha-latrotoxin receptor; CIRL; trkB; synapsin 1A;
 KW epithelial discoidin domain receptor 1; synapsin 1B; neuroleptic;
 KW tumour necrosis factor alpha; TNF-alpha; rat.
 XX
 OS Rattus sp.
 XX
 PN WO200175440-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 02-APR-2001; 2001WO-GB001486.
 XX
 PR 31-MAR-2000; 2000GB-00007880.
 PR 26-MAY-2000; 2000GB-00012768.
 XX
 PA (WELF-) WELFIDE CORP.

XX
 PI Cochran S, Paterson G, Ohashi Y, Morris B, Pratt J;
 XX WPI; 2002-010813/01.
 DR N-PSDB; AAS16843.
 XX
 PT Novel chronic animal model of schizophrenia, useful for identifying anti-
 PT psychotic drugs and genes that are associated with schizophrenia.
 XX
 PS Disclosure; Fig 8b; 79pp; English.
 XX
 CC The invention relates to YSG polynucleotide fragments for use in
 CC diagnosing and/or developing treatments for schizophrenia using chronic
 CC animal models. The polynucleotides and their encoded polypeptides are
 CC used for identification of compounds which modulate the expression of YSG
 CC molecules, leading to the manufacture of schizophrenia medications. The
 CC sequences can also be used for testing candidate compounds for any effect
 CC on the polypeptides. Anti-schizophrenic effects of a compound can be
 CC determined by measuring local cerebral glucose utilisation (LCGU) or
 CC comparing its expression level with that of a control group. The
 CC sequences are useful in the identification of genes associated with
 CC schizophrenic states and in the development of an antibody. The sequences
 CC of the invention include phosphodiesterase 1-alpha, calcium-independent
 CC alpha-latrotoxin receptors (CIRL)-1,2&3 epithelial discoidin domain
 CC receptor 1 (trkB), netrin receptor (UNC5H1), synapsins 1A and 1B and
 CC tumour necrosis factor (TNF) alpha. This sequence represents rat netrin
 CC receptor UNC5H1 (YSG7) polypeptide
 XX
 SQ Sequence 898 AA;
 Query Match 95.0%; Score 4545.5; DB 5; Length 898;
 Best Local Similarity 94.7%; Pred. No. 0;
 Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;
 QY 1 MAVRPGLPALLGLVLAALRGSGAQOSATVPANPGANPDLLPHFLVEDEVDVIVKNKP 60
 Db |||||
 QY 1 MAVRPGLPVLLGLVLAALRGSGAQOSATVPANPGANPDLLPHFLVEDEVDVIVKNKP 60
 Db |||||
 QY 61 VLLVCKAVPATQIFKCNGBWVRQVDHVIERSDSSGSEPTMEVIRINVSQOQVKVFGLE 120
 Db |||||
 QY 61 VLLVCKAVPATQIFKCNGBWVRQVDHVIERSDSSGSEPTMEVIRINVSQOQVKVFGLE 120
 Db |||||
 QY 121 EYWCOCVAMSSGTTKQKAYIRIARLKNPQEPPLAKEVSLQGVILPCRPPIGPAE 180
 Db |||||
 QY 121 EYWCOCVAMSSGTTKQKAYIRIARLKNPQEPPLAKEVSLQGVILPCRPPIGPAE 180
 Db |||||
 QY 181 VEMLRNEDLVDPSPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSAAVIVY 240
 Db |||||
 QY 181 VEMLRNEDLVDPSPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSAAVIVY 240
 Db |||||
 QY 241 VNGWSTWTEWVSCASCGRGWKRSRSCNTPAPLNGAFCEGQNV-QKTACATLCPVDG 300
 Db |||||
 QY 241 VNGWSTWTEWVSCASCGRGWKRSRSCNTPAPLNGAFCEGQNV-QKTACATLCPVDG 299
 Db |||||
 QY 301 SWSPKWSACGLDCTHWRSECDPAPRNGEBCQGTDLDRNCTSDLCVHSASGPEV 360
 Db |||||
 QY 300 SWSPKWSACGLDCTHWRSECDPAPRNGEBCGADLDRNCTSDLCVHSASGPEV 359
 Db |||||
 QY 361 ALYVGLIAVAVCLVLLVLLVYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL 420
 Db |||||
 QY 360 ALYIGLVAVAVCLFLLLLALGLIYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL 419
 Db |||||
 QY 421 LTIQPDLS-TTTTYQGSCLPRQDGPSPKQFQLTNGHLLSPLGGGRHTLHSSPTSEAEFV 479
 Db |||||
 QY 420 LTIQPDLS-TTTTYQGSCLPRQDGPSPKQFQLTNGHLLSPLGGGRHTLHSSPTSEAEFV 479
 Db |||||
 QY 480 SRLSTONYFRSLPRGTSNNYGTFTNFGRLMIPNTGISLLIPDAIPRGKIYEIYTLH 539
 Db |||||
 QY 480 SRLSTONYFRSLPRGTSNNYGTFTNFGRLMIPNTGISLLIPDAIPRGKIYEIYTLH 539
 Db |||||
 QY 540 KPDEVRLPLAGCQTLISPIVSCGPGVLLTRPVILAMDHCGEPSPDSWSLRKKQSCGS 599
 Db |||||
 QY 540 KPDEVRLPLAGCQTLISPIVSCGPGVLLTRPVILAMDHCGEPSPDSWSLRKKQSCGS 599
 Db |||||

```
QY 600 WEQDVHLHGEAPSHLYYCOLEASACVYFTTEOLGRFALVGEALSVAAAKRLKLLFPAPVA 659
Db 600 WE-DVLHGEESPSPHLYYCOLEAGACVYFTTEOLGRFALVGEALSVAAATKRLRLFPAPVA 658
QY 660 CTSLEYNIRVYCLHDTDALKEVVOLKLGQQLIQEPRVLHFKDSYHNLRSLSHDVPSS 719
Db 659 CTSLEYNIRVYCLHDTDALKEVVOLKLGQQLIQEPRVLHFKDSYHNLRSLSHDVPSS 718
QY 720 LMKSKLLVSVQEIFYHWNQYRHLCTFTLERVSPSTSDLACKLVMWQVGGDQSFSI 779
Db 719 LMKSKLLVSVQEIFYHWNQYRHLCTFTLERINASTSDLACKLVMWQVGGDQSFNI 778
QY 780 NFNITKTTRFAELLALSEAGVPALVGSAPFKIPFLIROKLIISSLDPPCRGGADWRTLAQ 839
Db 779 NFNITKTTRFAELLALSEGGVPALVGSAPFKIPFLIROKLIIASLDPPCSRGGADWRTLAQ 838
QY 840 KLHLDSHLSPFASKPSPMTAMTLNWEARHPFNGNLSQLAAVAGLQPDAGLFTVSEABEC 899
Db 839 KLHLDSHLSPFASKPSPMTAMTLNWEARHPFNGNLSQLAAVAGLQPDAGLFTVSEABEC 898

RESULT 6
AAU97900
ID AAU97900 standard; protein; 898 AA.
XX
AC AAU97900;
XX
DT 27-AUG-2002 (first entry)
XX
DE Rat netrin binding membrane receptor UNC5H-1 protein.
XX
KW Netrin binding membrane receptor; receptor; UNC5H-1; Rat; neurotropic;
KW neuroprotective; cytosolic; antiparkinsonian; cerebroprotective; cancer;
KW central nervous system; CNS; stroke; Parkinson's disease;
KW multiple sclerosis; Alzheimer's disease.
XX
OS Rattus sp.
XX
Key Location/Qualifiers
FH Domain 152..223
FT Domain /note= "Immunoglobulin domain "
FT Domain 247..294
FT Domain /note= "Thrombospondine type 1 domain "
FT Domain 302..348
FT Region /note= "Thrombospondine type 1 domain"
FT FT 361..382
FT FT /note= "Transmembrane region"
FT FT 495..598
FT FT /note= "ZUS domain"
FT FT 817..897
FT FT /note= "Death domain"
XX
PN WO200233080-A2.
XX
PD 25-APR-2002.
XX
PF 15-OCT-2001; 2001WO-EP011891.
XX
PR 16-OCT-2000; 2000US-0240061P.
XX
PA (FARB ) BAYER AG.
XX
PI Koehler RH;
XX
WPI; 2002-463314/49.
XX
PT Novel human netrin binding membrane receptor polypeptide and
PT polynucleotides for identifying modulating agents useful in treating
PT diseases e.g. Parkinson's disease, multiple sclerosis, stroke,
PT Alzheimer's disease.
XX
PS Disclosure; Fig 3; 94pp; English.
```

```
XX This invention relates to the DNA and protein sequences of a novel
CC purified human netrin binding membrane receptor, UNC5H-1. The DNA
CC sequence of the invention is useful as a probe for detecting a nucleic
CC acid encoding the UNC5H-1 protein in a biological sample. The sequences
CC of the invention are useful to screen for agents which decrease the
CC activity of the UNC5H-1 protein. The sequences are also useful for
CC screening agents which regulate (modulate) the activity of the protein of
CC the invention. A pharmaceutical composition containing the protein of the
CC invention or a reagent that modulates the activity of the UNC5H-1 protein
CC may be useful for treating a UNC5H-1 dysfunction related disease such as
CC cancer or a central nervous system (CNS) disorders (e.g., Parkinson's
CC disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion
CC proteins comprising the UNC5H-1 protein are useful for generating
CC antibodies and for in various assay systems, and the protein can be used
CC as a bait protein in a two-hybrid assay or three-hybrid assay. The method
CC of the invention is useful for detecting a coding sequence for the UNC5H-
CC 1 protein. The present sequence represents the Rat netrin binding
CC membrane receptor UNC5H-1 protein of the invention
XX
SQ Sequence 898 AA;

Query Match 95.0%; Score 4545.5; DB 5; Length 898;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

QY 1 MAVRPGLPALLGIVLAALRGSGAQOSATVANPVGANPDLLPHFLVEPEDVIVKNKP 60
Db 1 MAVRPGLPVLLGIVLAALRGSGAQOSATVANPVGANPDLLPHFLVEPEDVIVKNKP 60
QY 61 VLLVCKAVPATQIFKCNKGEWVRQVDHVIERTSDGSSGEPTMEVRINVSROQVEKVFGL 120
Db 61 VLLVCKAVPATQIFKCNKGEWVRQVDHVIERTSDSSGLPTMEVRINVSROQVEKVFGL 120
QY 121 EYWCQCVAMSSGTTKQKAYIRIARLKNFEQPLAKEVSLKQGIPLCPPEGIPPAE 180
Db 121 EYWCQCVAMSSGTTKQKAYIRIARLKNFEQPLAKEVSLKQGIPLCPPEGIPPAE 180
QY 181 VEWLNRDLVDPDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVY 240
Db 181 VEWLNRDLVDPDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVY 240
QY 241 VNGGKSTWTSVCSASCGRWQKRSCTNPAPLNGGAFCEGQNVHDTVSSLLVSDG 300
Db 241 VNGGKSTWTSVCSASCGRWQKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299
QY 301 SWSWSKWSACGLDCTHWRSECDPAPRNGEGECQGTDLTRNCTSLCVHSASGPDV 360
Db 300 SWSWSKWSACGLDCTHWRSECDPAPRNGEGECRGADLDRNCTSLCVHTASCPDV 359
QY 361 ALYGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL 420
Db 360 ALYIGLVAVAVCLFLLLLALGLIYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL 419
QY 421 LTIQPDLS-ITTTTYGSLCPQDQSPKQFTNGHLLSLPLGGGRHTLHHSPTSABEFV 479
Db 420 LTIQPDLSITTTTYGSLCSRDQGPSPKFQLSNGHLLSLPLGGGRHTLHHSPTSABEFV 479
QY 480 SRLSTQNYFRSLPRGTSNMTYGTNFGRLMIPNTGISLLIPDAIPRGKIYEIYLTLH 539
Db 480 SRLSTQNYFRSLPRGTSNMTYGTNFGRLMIPNTGISLLIPDAIPRGKIYEIYLTLH 539
QY 540 KPEDVRLPLAGCOTLLSPVSCGPPGVLLTRPVILAMDHCGEPPSDSLRLKKQCEGS 599
Db 540 KPEDVRLPLAGCOTLLSPVSCGPPGVLLTRPVILAMDHCGEPPSDSLRLKKQCEGS 599
QY 600 WEQDVHLHGEAPSHLYYCOLEASACVYFTTEOLGRFALVGEALSVAAAKRLKLLFPAPVA 659
Db 600 WE-DVLHGEESPSPHLYYCOLEAGACVYFTTEOLGRFALVGEALSVAAATKRLRLFPAPVA 658
QY 660 CTSLEYNIRVYCLHDTDALKEVVOLKLGQQLIQEPRVLHFKDSYHNLRSLSHDVPSS 719
Db 659 CTSLEYNIRVYCLHDTDALKEVVOLKLGQQLIQEPRVLHFKDSYHNLRSLSHDVPSS 718
```

QY 720 LWSKLLVSYQEIPIFYHWMGTORYLHCTFTLERSVSPSTDLACKLWVWQVGGQSPSI 779
 Db 719 LWSKLLVSYQEIPIFYHWMGTORYLHCTFTLERSVSPSTDLACKLWVWQVGGQSPSI 778
 QY 780 NFNTKTRFAELLALSEAGVPALVGFSAFKIPFLIRQKIISLDPCCRRGADWRTLAQ 839
 Db 779 NFNTKTRFAELLALSEAGVPALVGFSAFKIPFLIRQKIISLDPCCRRGADWRTLAQ 838
 QY 840 KLHLDLSLSPFASKPSTAMILNWEARHPFNGNLSQLAAVAGLQDPDAGLFTVSEAC 899
 Db 839 KLHLDLSLSPFASKPSTAMILNWEARHPFNGNLSQLAAVAGLQDPDAGLFTVSEAC 898

RESULT 7
 AAM79128
 ID AAM79128 standard; protein; 943 AA.
 XX AC AAM79128;
 XX DT 06-NOV-2001 (first entry)
 XX DE Human protein SEQ ID NO 1790.
 XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX OS Homo sapiens.
 XX PN WO200157190-A2.
 XX PD 09-AUG-2001.
 XX PF 05-FEB-2001; 2001WO-US004098.
 XX PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI; 2001-476283/51.
 XX N-PSDB; AAK52261.
 XX PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 XX PS Claim 20; Page 4148-4150; 6221pp; English.
 XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication

XX SQ Sequence 943 AA;
 Query Match 92.6%; Score 4434; DB 4; Length 943;
 Best Local Similarity 90.4%; Pred. No. 0;
 Matches 853; Conservative 4; Mismatches 17; Indels 70; Gaps 7;
 QY 1 MAYRPGIWPALLGIVLAALWLRGSAQOQSATVAVPVGANPDLLPHFLVEPDEVIVNKP 60
 Db 25 MTRRPSL-----MAGRQHGWSAQOQSATVAVPVGANPDLLPHFLVEPDEVIVNKP 76
 QY 61 VLVVCKAVPATQIFFKCNCEWVRQDVHVIERSDGSSEPTMEVRINVSQOQVKVFGLE 120
 Db 77 VLVVCKAVPATQIFFKCNCEWVRQDVHVIERSDGSSEPTMEVRINVSQOQVKVFGLE 136
 QY 121 EYWCQCVMSSSGTTKQKAYIRIARLRNFEQPLAKEVLSLEGGIVLPCRPPEGIPAE 180
 Db 137 EYWCQCVMSSSGTTKQKAYIRIARLRNFEQPLAKEVLSLEGGIVLPCRPPEGIPAE 196
 QY 181 VEMLRNEDLVDPISLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIV 240
 Db 197 VEMLRNEDLVDPISLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIV 256
 QY 241 -----VNGCWSTWTSVCSASCGRCWKRSRSTN 271
 Db 257 GPRDSILVTGRGTAVPLGSDMWLSFVVRPVNGGWSWTWTSVCSASCGRCWKRSRSTN 316
 QY 272 PAPLNGAFCEGQNVHRTVSSLLVSDGSWSPWSKWSACGLDCTHWSRECSDDPAPRNG 331
 Db 317 PAPLNGAFCEGQNV-KTACATLCPVDGSWSPWSKWSACGLDCTHWSRECSDDPAPRNG 375
 QY 332 GEBCQGTDLDRNCTSDLCVH-----SASGPEDVALYVGLIAVAVCLVLL 376
 Db 376 GEBCQGTDLDRNCTSDLCVHNYTPTAKMLSPAAASGPEVALYVGLIAVAVCLVLL 435
 QY 377 LVLVLVYCKRKEGLSDVDADSSILTSFGFOPVSIKPSKADNPHLLITQPDLS-TTTTYQG 435
 Db 436 LVLVLVYCKRKEGLSDVDADSSILTSFGFOPVSIKPSKADNPHLLITQPDLS-TTTTYQG 495
 QY 436 SLCPRODGPSPKQLTNGHLLSPLGGRHTLHSSPTSEAEFVSRLSTQNYFRSLPRGT 495
 Db 496 SLCPRODGPSPKQLTNGHLLSPLGGRHTLHSSPTSEAEFVSRLSTQNYFRSLPRGT 555
 QY 496 SNMTYGTFFLGGRLMPNTGISLIPDPAIPRGKIYIYLTLHKPDVRLPLAGCOTLL 555
 Db 556 SNMTYGTFFLGGRLMPNTGISLIPDPAIPRGKIYIYLTLHKPDVRLPLAGCOTLL 603
 QY 556 SPIVSCGPPGVLLTRPVILAMDHCGEPPSDSNLRLKQSCGSWEODVLHLGEEAPSHL 615
 Db 604 ---VSCGPPGVLLTRPVILAMDHCGEPPSDSNLRLKQSCGSWE-DVLHLGEEAPSHL 659
 QY 616 YYQLEASACYVTEQGRFALVGEALSAVAAAKRLKLLFPAPVACTSLEYNIRVYCLHDT 675
 Db 660 YYQLEASACYVTEQGRFALVGEALSAVAAAKRLKLLFPAPVACTSLEYNIRVYCLHDT 719
 QY 676 HDALKEVQLEKQGLQLOEPRVLFHFKDSYHNLRLSIHDVPSLLKSKLLSYQIPFY 735
 Db 720 HDALKEVQLEKQGLQLOEPRVLFHFKDSYHNLRLSIHDVPSLLKSKLLSYQIPFY 779
 QY 736 HIWNGTQRYLHCTFTLERSVSPSTDLACKLWVWQVGGQSPSINFTKTRFAELLAL 795
 Db 780 HIWNGTQRYLHCTFTLERSVSPSTDLACKLWVWQVGGQSPSINFTKTRFAELLAL 839
 QY 796 ESEAGVPALVGFSAFKIPFLIRQKIISLDPCCRRGADWRTLAQKHLDSLHSLFFASKPS 855
 Db 840 ESEAGVPALVGFSAFKIPFLIRQKIISLDPCCRRGADWRTLAQKHLDSLHSLHSLFFASKPS 899
 QY 856 PTAMILNLWEARHPFNGNLSQLAAVAGLQDPDAGLFTVSEAC 899
 Db 900 PTAMILNLWEARHPFNGNLSQLAAVAGLQDPDAGLFTVSEAC 943

AAU74818
 ID AAU74818 standard; protein; 842 AA.
 AC AAU74818;
 XX
 XX 23-APR-2002 (first entry)
 DT
 XX Human REPTR 1 protein.
 DE
 XX
 XX REPTR; human; antiinflammatory; cytostatic; immunosuppressive; antiviral;
 KW anti-HIV; antiarthritic; anticonvulsant; nootropic; neuroprotective;
 KW antiallergic; antibody; immunogen; endometriosis;
 KW gastrointestinal disorder; gastritis; oesophageal carcinoma;
 KW Crohn's disease; irritable bowel syndrome; ulcerative colitis;
 KW endocrine disorder; hypothalamus disorder; Kallman's disease;
 KW autoimmune disease; inflammatory disease; infertility; receptor;
 KW acquired immune deficiency syndrome; AIDS; rheumatoid arthritis; allergy;
 KW osteoarthritis; diabetes mellitus; multiple sclerosis;
 KW systemic lupus erythematosus; cell proliferative disorder; cancer;
 KW developmental disorder; Duchenne muscular dystrophy;
 KW Becker muscular dystrophy; neurological disorder; epilepsy;
 KW Alzheimer's disease; Huntington's disease; reproductive disorder.
 XX
 OS Homo sapiens.
 XX
 XX WO200198354-A2.
 PN
 XX 27-DEC-2001.
 PD
 XX
 XX 21-JUN-2001; 2001WO-US019942.
 PF
 XX
 XX 21-JUN-2000; 2000US-0214027P.
 PR
 XX 25-AUG-2000; 2000US-0228045P.
 PR
 XX 12-DEC-2000; 2000US-0255104P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Griffin JA, Kallick DA, Tribouley CM, Yue H, Nguyen DB, Tang YT;
 PI Lal P, Policky JL, Azimzai Y, Lu DAM, Graul R, Yao MG, Burford N;
 PI Hafalia AJA, Baughn MR, Bandman O, Patterson C, Yang J, Xu Y;
 PI Gandhi AB, Warren BA, Ding L, Sanjanwala MS, Duggan BM, Lu Y;
 XX WPI: 2002-090432/12.
 DR
 DR N-PSDB; ABK35169.
 XX
 XX Twelve human receptors (referred to as REPTR-1 to REPTR-12), useful in
 PT the diagnosis, treatment and prevention of gastrointestinal (e.g.
 PT gastritis), autoimmune/inflammatory (e.g. osteoarthritis) and cell
 PT proliferative (e.g. cancer) disorders.
 XX
 XX Claim 45; Page 111-113; 157pp; English.
 PS
 XX This invention relates to twelve human receptors cDNA sequences referred
 CC to as REPTR-1 to REPTR-12), and the proteins encoded thereby. The
 CC proteins of the invention may have antiinflammatory, cytostatic,
 CC immunosuppressive, antiviral, anti-HIV, antiarthritic, muscular active
 CC general, anticonvulsant, nootropic, neuroprotective, antiallergic
 CC activities. The sequences of the invention may be used to produce REPTR
 CC agonists or antagonists, and the protein sequences may be used to raise
 CC anti-REPTR antibodies. These molecules and the REPTR polynucleotides and
 CC polypeptides of the invention are useful in the diagnosis, treatment and
 CC prevention of gastrointestinal (e.g. gastritis, oesophageal carcinoma,
 CC Crohn's disease, irritable bowel syndrome, ulcerative colitis), endocrine
 CC (e.g. hypothalamus disorder, Kallman's disease), autoimmune/ inflammatory
 CC (e.g. acquired immune deficiency syndrome (AIDS), rheumatoid arthritis,
 CC allergies, osteoarthritis, diabetes mellitus, multiple sclerosis,
 CC systemic lupus erythematosus), cell proliferative (e.g. cancer),
 CC developmental (e.g. Duchenne and Becker muscular dystrophy), neurological
 CC (e.g. epilepsy, Alzheimer's disease, Huntington's disease) and
 CC reproductive (e.g. infertility, endometriosis) disorders. Numerous other
 CC examples of each disorder are given in the specification. The present
 CC sequence represents the human REPTR1 protein sequence of the invention
 XX

SQ Sequence 842 AA;
 Query Match 91.5%; Score 4379.5; DB 5; Length 842;
 Best Local Similarity 93.1%; Pred. No. 0;
 Matches 838; Conservative 1; Mismatches 2; Indels 59; Gaps 3;
 QY 1 MAVRPGIWPALLGIVLAAWLRGSGAQOSATVANPVPGANPDLLPHFIVEDEVIVVKNKP 60
 DB 1 MAVRPGIWPALLGIVLAAWLRGSGAQOSATVANPVPGANPDLLPHFIVEDEVIVVKNKP 60
 QY 61 VLVCKAVPATQIFFKCNCEWVRQVDHVIERSDSSGSEPTMEVRINVSQQVEKVGLE 120
 DB 61 VLVCKAVPATQIFFKCNCEWVRQVDHVIERSDSSGSEPTMEVRINVSQQVEKVGLE 120
 QY 121 EYWCQCVAMSSGTTKSKAYIRIARLRKPEPEPLAKEVLSBOGIVLPCRPPEGIPPAE 180
 DB 121 EYWCQCVAMSSGTTKSKAYIRIARLRKPEPEPLAKEVLSBOGIVLPCRPPEGIPPAE 180
 QY 181 VEWLNRNEDLVDPNVIITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
 DB 181 VEWLNRNEDLVDPNVIITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
 QY 241 VNGCWSTWTEWSVCSASCGRGWQKRSRSCNPNAPLNGGAFCEGQNVHRTVSSLLVSDG 300
 DB 241 -----VDG 243
 QY 301 SWSPWKSWSACGLDCTHWSRECSDDPAPRNGBECQGTDLDRNCTSDLCVHSASGEDV 360
 DB 244 SWSPWKSWSACGLDCTHWSRECSDDPAPRNGBECQGTDLDRNCTSDLCVHTASGEDV 303
 QY 361 ALYVGLIAVAVCLVLLVLLVLYVCRKKEGLSDSDVADSSILTSQFQVSKADNPHL 420
 DB 304 ALYVGLIAVAVCLVLLVLLVLYVCRKKEGLSDSDVADSSILTSQFQVSKADNPHL 363
 QY 421 LTIQPDLS-TTTTYQSGSLCPRQDGPSPKFQLTNGHLLSPGGGRHTLHSSPTSEAEFFV 479
 DB 364 LTIQPDLS-TTTTYQSGSLCPRQDGPSPKFQLTNGHLLSPGGGRHTLHSSPTSEAEFFV 423
 QY 480 SRLSTQNYPSRLPRGTSNMTYTFNFGGLRMLPNTGISLLIPDAIPRGKIYIYTLH 539
 DB 424 SRLSTQNYPSRLPRGTSNMTYTFNFGGLRMLPNTGISLLIPDAIPRGKIYIYTLH 483
 QY 540 KPEDVRLPLAGCOTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLLKQSCGS 599
 DB 484 KPEDVRLPLAGCOTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLLKQSCGS 543
 QY 600 WEQDVLHGBEAPSHLYYCOLEASACVYFTQGRFALVGEALSVAALKLLLEAPVA 659
 DB 544 WE-DVLHGBEAPSHLYYCOLEASACVYFTQGRFALVGEALSVAALKLLLEAPVA 602
 QY 660 CTSLEYNIRVYCLHDTHDALKEVVOLEKQGGQLIQEPVRLHFKDSYHNRLSLTHDVPSS 719
 DB 603 CTSLEYNIRVYCLHDTHDALKEVVOLEKQGGQLIQEPVRLHFKDSYHNRLSLTHDVPSS 662
 QY 720 LMKSKLLVSVQEPFYHIWNGTQRYLHCTFTLERSVSPSTDACKLWVWQVEGQGSFSI 779
 DB 663 LMKSKLLVSVQEPFYHIWNGTQRYLHCTFTLERSVSPSTDACKLWVWQVEGQGSFSI 722
 QY 780 NFNITKDTREFAELIALESEAGVPALVGPSAFKIPFLIRQKIIISLDDPCRRGADWRTLAQ 839
 DB 723 NFNITKDTREFAELIALESEAGVPALVGPSAFKIPFLIRQKIIISLDDPCRRGADWRTLAQ 782
 QY 840 KLHLDSHLSFASKPSPMTAMLNWEARHPNGNLSOLAARVAGLGOPDAGLFTVSEAC 899
 DB 783 KLHLDSHLSFASKPSPMTAMLNWEARHPNGNLSOLAARVAGLGOPDAGLFTVSEAC 842
 RESULT 9
 AAU78899
 ID AAU78899 standard; protein; 556 AA.
 XX
 XX AAU78899;
 XX

PT identifying unknown human cDNAs which encode proteins that interact with
PT the human unc-5C protein.

PT the human unc-5C protein.

PS Disclosure: Page 224-227; 246pp; English.

The present invention describes 3 variants of human unc-5C cDNAs (unc-5Cb, unc-5Cc and unc-5C8) which correspond to alternatively spliced unc-5C transcripts, and a human unc-5Hs1 cDNA which shares homology with the *Rattus norvegicus* unc-5Hs1 cDNA. Also described are assays based on protein-protein-interactions between the unc-5 protein and a variety of different protein-interacting proteins. The unc-5C variant cDNAs and unc-5Hs1 cDNA are useful in methods for identifying compounds which reduce or inhibit the lethal phenotype associated with the expression of the unc-5 death domain in yeast. They are also useful in yeast two hybrid experiments for identifying unknown human cDNAs which encode proteins that interact with the human unc-5C protein. AAC90914 to AAC90971 and AAB50646 to AAB50693 represent sequences used in the exemplification of the present invention.

Sequence 931 AA;

Query Match

Query Match	Score 2673.5; DB 4
55.8%;	

Best Local Similarity	55.5%;	Pred. No. 1.1e-222;	
Matches	507;	Conservative	154; Mismatches 221; Indels 31; Gaps 12;

QY 9 PALLGIVLAAWLRGSGAQQS---ATVANPVGANPDLLPHFLVEPEDVYVKNKPVLLVC 65

26 PAT - ATISACTCSAONDDFEHET.PETERSDDPERI.PHEI.TEPFAYVKNKPNT.YC 83

QY 66 KAVPATQIFFKNCGEWVRQVDHVIERTDSSGEPIMEVRINVSRRQVVKVFGLEEYWCQ 125

db 84 KASPA^TLYFKCNSEVHOKDHVDERVDETSGLIVREVSIEISRQVVEELFGPEDIWQC 143

Qy 126 CVAMSSGTTKSQAYIRIARLRLKNFQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLR 185

Db 144 CVAWSAGTTKSRKAYVRIAYIRKTFEQEPLGKEVSLQCRPPPEGIPVAEVEWLK 203

Qy 186 NEDLVPSLDPNVYITREHSLVVRQARLADTANTYTCVAKNIVARRRSASAAVIVYVNGW 245

Db 204 NEDIIDPVEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTATVIVYNGW 263

[illegible]

Db 264 STWTENSVNSRCGRGYQKRTICTNPAPLNGGAFCEQSV-QKIACTILCPVDGRWTPW 322

[illegible]

D6 323 SKWSTCGTCTHWRRRRECTAPAPKNGGKCDGLVLQSNCTDGLCMQTAPSDJVALYVG 382

[illegible][illegible]

db 440 PDLTSAAMVRGPPVYALHD-VSDKIPMTNSPILDPNLKIKVYNTSGAVSPQDDLSEFT 498

480 SPI S---TONYE-----RSIPBGT--SNMTYGTENFEGGRIMIPNTGTSIUJPPDA 525

Db 499 SKLSPQWTOILLENEALSXNQSLARQTDPSCTAFGSFNSLGGHLIVPNSGVSLIPAGA 558

526 IPRGKIYEIYLTJHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGESPD 585

Db 559 IPQGRVYEMYYVHRKETMRPPMDDSQTLTPWSCPPGALLTRPVWLTMHHCADPNT 618

QY 586 SWSLRLKKQSCGSEWDVLHLGEEAPSHLYYCQLEASACVFTTEQLGRFALVGEALSA 645

Db 619 DWKILLKNOAQCQWE-DVVVVGEENFTPCYIKLDAEACHILTENLSTYALVGHSTTKA 677

QY 646 AAKRLKLLFAPVACTSLEYNIRVYCLHDTHDALKEVVLEKQLGGQLIQEPRVLHFKDS 705

678 AAKRLKLAIFGPLCCSSLEYSIRVICLDLITQDALKEILHLEKQIGGGDEEFPALHFGS 73

CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 931 AA;

Query Match 55.8%; Score 2673.5; DB 7; Length 931;
 Best Local Similarity 55.5%; Pred. No. 1.1e-222;
 Matches 507; Conservative 154; Mismatches 221; Indels 31; Gaps 12;

Qy 9 PALLGIVLAWLRSQAGQSS---ATVANPVGANPDLLPHFLVPEPDYVIVKNKPVLLVC 65
 Db 26 PAL--ALLSASGTSAAGQDDFFHELPEPTSPDPPEPLHELPEPEAYIVKNKPVNLYC 83
 Qy 66 KAVPATQIFKCGEWRQVDHVIERSDSSGSEPTMEVRINVRQVEKVGLEEWYCQ 125
 Db 84 KASPATQIFYKCGEWRHQDHIIVDERVDETSGLIVREVSEISRQVVEELFGPEDEWYCQ 143
 Qy 126 CVAWSSGTTKSKQAYIRIARLRKNFEQPLAKEVLEQGIPLCRPPEGIPPAEVEWLR 185
 Db 144 CVAWSSAGTTKSKAYVRIARVIRTFEPEPLGKEVSELEQVLLQCRPPEGIPPAEVEWLR 203
 Qy 186 NEDLVPSLDPNVITREHSLVVRQARLADPANTVCVAKNIVARRRSASAAVIVVNGW 245
 Db 204 NEDIIDEVDENFYITIDHNLIIKQARLSDTANTVCVAKNIVAKRKSTTATVIVVNGW 263
 Qy 246 STTWSVCSASCGRGWQKSRCTNPAPNGAFCEGQNVHDTVSSLLVSDGSPW 305
 Db 264 STTWSVCSNRGQYQKTRCTNPAPNGAFCEGQSV-QKIACTTCLCPVDGRWTPW 322
 Qy 306 SKWSACGLDCTHWSRECSPPAPNGGEECGTDLDRNCTSDLCVHSASGPEVVALYVG 365
 Db 323 SKWSTCGTECHWRRECTAPAPNGGKDCDGLVQSKNCTDGLCMQTAPDSDVVALYVG 382
 Qy 366 L-TAVAVCLVLLLVILVYCRKKGDSVDVADSSILTSFGFQVSIKPSKADNPHELLTIQ 424
 Db 383 IIVAVIVCLAVVVALFVYKKNHRDPESDIISSALNGGFQVNIKAARQD---LLAVP 439
 Qy 425 PDL-STTTTYGSLCPQDQSPKFOITNGHLLSPGGRHLLHS-----SPTSEAEFV 479
 Db 440 PDLTSAAMRYGPIVALHD-VSDKIPMTNBPILDPLPNLKIYNTSGAVSPQDDLSEFT 498
 Qy 480 SRLS---TQNYF-----RSLPRGT--SNMTYGTFFLGGRLMIPNTGISILLIPDA 525
 Db 499 SKLSPQMTQSLLENEALS LKXNSLARQTDPSCTAFGSFNSLGHILVFNISGVLLIPAGA 558
 Qy 526 IPKGKYEIVLTHKPDVRLPLAGCOTLLSPVSCGPPGVLLTRPVILAMDHCGEFPD 585
 Db 559 IPOGRVYEMVTYVHRKETWAPMDSDSQTLLTPVVVSCGPGALLTRPVILTMHCADNTE 618
 Qy 586 SWSLRLKQSCGSEWDLHLGEEAPSHLYYQLEASACVYVTEQGRFALYVGEALSVA 645
 Db 619 DWKILLKNAQAQGW-E-DVVVGEENFTTFCYKLDAAECHILTENLSTVALYVGHSTTKA 677
 Qy 646 AAKRLKLLPAPVACTSLEYNIRVYCLDTHDALKEVVQLEKOLGGOLIOEPRLVHFKDS 705
 Db 678 AAKRLKLAIFGLCCSLEYSIRVYCLDTHDALKEILHLERQTGGQLLEEPKALHFKGS 737
 Qy 706 YHNLRLSHDVPSSLKMSKLLVSYQEIPIFYHWNQYRVLHCTFTLERSVPSSTDLACKL 765
 Db 738 THNLRLSIHDIAHSLMWSKLLAKYQEIPIFYVHWSGQRNLHCTFTLERSVNLVELVCKL 797
 Qy 766 WYMQVEGDSGFSINFNITKDRFAELLALLESAGVPALVGPFAKIPPIRQKILSSLD 825
 Db 798 CVRQVEGEGQIFQINCTVSEPTGIDLPDLPDANTITVTGSPAFSIPPLRQKLCSSLD 857

Qy 826 PPCRRGADWRTLAOKLHLDLSLSPFASKPSPPTAMILNLWEARHPNGNLSQLAARAVAGLG 885
 Db 858 APQTRGHDWRMLAKHLNLDRLYNLFATKSSPTGVILDWEAQNPDPGNLSMLAAVLEEMG 917
 Qy 886 QPDAGLFTVSEAE 898
 Db 918 RHETVVSIAAEGQ 930

RESULT 12

ABG11551
 ID ABG11551 standard; protein; 982 AA.

XX AC ABG11551;
 XX DT 18-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #11542.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDE; AAS75738.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 20; SEQ ID NO 41910; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 982 AA;

Query Match 55.8%; Score 2673.5; DB 4; Length 982;
 Best Local Similarity 55.5%; Pred. No. 1.3e-222;

Matches 507; Conservative 154; Mismatches 221; Indels 31; Gaps 12;	
Qy	9 PALGIVLAWLKRGGAQGS---ATVANPVGANPDLPHLPVPEPDIYVKNKPVLLVC 65
Db	77 PAL--ALLSAGTGSAAQDDDFHELPEPDPPEPLPFLIBPEEAYIVKNKPVNLYC 134
Qy	66 KAVPATQIFKNCGEWVRQVDHVIERTDGGSGEPTMEVRINVSRQVQKVFGLVEEYWCQ 125
Db	135 KASPATQIFKNCSEWVHQDHIHIVRDETSGLIVREVSIEISRQVQVEELFGEDYWCQ 194
Qy	126 CVAMSSGTTKSKAYIRIARLRNFRQEPLEAKESVLEQGVILPCRPPEGIPPAEVEWLR 185
Db	195 CVAMSSAGTTKSKAYVRIAYLRKTFEQLPGLKEVSLQEVLLQCRPEGIPVAEVEWLK 254
Qy	186 NEDLVDFSLDPNVITREHSLVVRQARLADTANTYCVAKNIVARRRSASAAVIYVNGGW 245
Db	255 NEDIIDPEVDRNFYITDHNLIIRKQARLSDTANTYCVAKNIVARRKSTTATVIYVNGGW 314
Qy	246 STWTWESVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVHVRTVSSLLVSVDSGNSPW 305
Db	315 STWTWESVCSNRGCGYQKTRTCTNTPAPLNGGAFCEGQSV-QKIACITLCPVDGRWTPW 373
Qy	306 SKMSACGLDCTHMRSRCSPPAPRNGGEECGTDLTRNCTSDLCVHSASGEFDVALYVG 365
Db	374 SKWSTGCTECHWRRRECTAPAKNGGKDCDGLVLQSKNCTDGLCMQTAPDSDVALYVG 433
Qy	366 L-TAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSQPVSIKPSKADNPHLLTIQ 424
Db	434 IVIAVIVCLAISSVVVALFVYRKNRDFESDIIDSSALNGGFPQVNIKAARQD---LLAVP 490
Qy	425 PDL-STTTTGGSLCPQDGPSPFQLTNGHLLPLGGGRTHHS---SPTSEAEFEV 479
Db	491 PDLTSAAMVKGVPYALHD-VSDKIPNTPSILPDLNPKIKVYNTSGANSPQDDLEFT 549
Qy	480 SRLS---TONYF-----RSLPRGT--SNMTYGTNFGRLMIPNTGISLLIPDA 525
Db	550 SKLSPQMTQSLLENEALSQKQSLARQTDPSCTAFGFSNLSGGLHIVNSGVSLIIPAGA 609
Qy	526 IPQKVIYIYLTKHPEDVRPLAGCQTLLSPVSCGPPGVLLTRPVILAMDHGEPSPD 585
Db	610 IPQGRVYEMVYVHRKETMRPMDSDQTLTFVWSCGPPGALLTRPVVLTWHHCADPNT 669
Qy	586 SWSRLKQKQSGEWEQDVLHGLGEAPSHLYVCOLEASACVFTQGLRFPALVGEALSVA 645
Db	670 DWKILLKQAAQGGWE-DVVVVGEEFTPCYIKLDAEACHILLENSTYALVGHSTYKA 728
Qy	646 AAKRLKLLFAPVACTSLEYNIRVYCLDTHDALKEVVQLKQLGGQLIQPRVLHFKDS 705
Db	729 AAKRLKLAIFGLPCCSSLEYSIRVYCLDQDQALKEILHLERQGTGGQLLEBPKALHPKGS 788
Qy	706 YHNLRLSIHDVPSSLWKSLLVSVQEIIPPYHIWNGTQRYLHCTFTLERSVSPSTDACKL 765
Db	789 THNLRLSIHDTHSLWKSLLAKYQEIIPFYHWGSGQSNLHCTFTLERSVSNLTVELVCKL 848
Qy	766 WWOVEGDGQGSFNFTIKDTREAEILALESEAGVPALGVPSAPKIPFLRQKIISLD 825
Db	849 CVRQVEGGQFQNLCTVSEPTGIDPLLPANTITVTGSPAFSIPDLPRQKLCSSLD 908
Qy	826 PPCRGADWRTLAQKHLDSHLSFASKSPSPMTAMLNLEWHPNGLNSQIAAAVAGLG 885
Db	909 APQTRGHDRWMLAHKLNLDRLYNLFATKSSPTGVILDLWEAQNFDPDGNLSMLAAVLEBMG 968
Qy	886 QPDAGLFTVSEAE 898
Db	969 RHETVVSIAEGQ 981

RESULT 13
ADE63096
ID ADE63096 standard; protein; 945 AA.
XX
AC ADE63096;
XX

DT	29-JAN-2004 (first entry)
XX	Rat Protein AAB57679, SEQ ID NO 9031.
XX	Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW	chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX	Rattus norvegicus.
OS	WO2003016475-A2.
PN	27-FEB-2003.
PD	14-AUG-2002; 2002WO-US025765.
XX	14-AUG-2001; 2001US-0312147P.
PR	01-NOV-2001; 2001US-0346382P.
PR	26-NOV-2001; 2001US-033347P.
XX	(GEHO) GEN HOSPITAL CORP.
PA	(FARB) BAYER AG.
XX	Woolf C, D'urso D, Befort K, Costigan M;
PI	WPI; 2003-268312/26.
XX	GENBANK; AAB57679.
DR	
DR	
XX	New composition comprising two or more isolated polypeptides, useful for
PT	preparing a medicament for treating pain in an animal.
PT	
XX	Claim 1; Page; 1017pp; English.
PS	
XX	The invention discloses a composition comprising two or more isolated rat
CC	or human polynucleotides or a polynucleotide which represents a fragment,
CC	derivative or allelic variation of the nucleic acid sequence. Also
CC	claimed are a vector comprising the novel polynucleotide, a host cell
CC	comprising the vector, a method for identifying a nucleotide sequence
CC	which is differentially regulated in an animal subjected to pain and a
CC	kit to perform the method, an array, a method for identifying an agent
CC	that increases or decreases the expression of the polynucleotide sequence
CC	that is differentially expressed in neuronal tissue of a first animal
CC	subjected to pain, a method for identifying a compound which regulates
CC	the expression of a polynucleotide sequence which is differentially
CC	expressed in an animal subjected to pain, a method for identifying a
CC	compound that regulates the activity of one or more of the
CC	polynucleotides, a method for producing a pharmaceutical composition, a
CC	method for identifying a compound or small molecule that regulates the
CC	activity in an animal of one or more of the polypeptides given in the
CC	specification, a method for identifying a compound useful in treating
CC	pain and a pharmaceutical composition comprising the one or more
CC	polypeptides or their antibodies. The polynucleotide or the compound that
CC	modulates its activity is useful for preparing a medicament for treating
CC	pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC	injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC	therapy). The sequence presented is a rat protein (shown in Table 2 of
CC	the specification) which is differentially expressed during pain. Note:
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic form directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences.
XX	Sequence 945 AA;
SQ	

Query Match 52.2%; Score 2497; DB 7; Length 945;	
Best Local Similarity 52.0%; Pred. No. 2.6e-207;	
Matches 501; Conservative 144; Mismatches 236; Indels 82; Gaps 20;	
Qy	1 MAVRPGMLPALLGIVLAAW-----LRG--SGAQSATVANPVGANPDLPHLPVPEBV 53
Db	1 MRASGARGALLLALLLCWDPTPSLAGIDSGQ---ALPDGFPGAPAPQLPHFLLEPEDA 57
Qy	54 YIVKNKPVLLVCKAVPATQIFKNCGEWVRQVDHVIERTDGGSGEPTMEVRINVSRQV 113
Db	58 YIVKNKPVLELHCRAPPATQIYFKNCGEWVSQKHVQTQESLDEATGLRIRREIQVIEVSQV 117

Best Local Similarity 52.1%; Pred. No. 9.7e-207;		Matches 496; Conservative 150; Mismatches 247; Indels 59; Gaps 18;	
QY	1	MAVRPGLWPAALLGIVLAAM---	LRSGAQAQATV-ANPVCANPDLPLPHELVPEPDIYV 56
Db	1	MGARSGARGALLALLCWDPRLSOAGRKSGEVLDPSPAPAEPLPYFLQEPQDAYIV 60	
QY	57	KNKPVLLVCKAVPATQIFPKNGHWVQVDHVIETSDGSS-----CEPTMEVRINVSQ 111	
Db	61	KNKPEVLRCAFFPATQIFYKNGEWSQNDHVTOQGLDEATLGARGGLRVREVOIEVSRQ 120	
QY	112	QVEKVFGLVEWCCVAMSSGTTKSKAYTRIALRKNPEOEPLAKEVSLQOQVILPCR 171	
Db	121	QVEELFGLEDYWCOCVAMSSAGITKSRAYVRIACLRKNPDQEPKGEVLDEHVEVLQCR 180	
QY	172	PPEGIPPAEVEWLRNEDLVPSLDPNYIITREHSLVVRQARLADANTYCVAKNIVARR 231	
Db	181	PPEGVPAEVEWLKNEVDIPTQDTNFTLLTIIDHNLIIIRQARLSDTANTYCVAKNIVARR 240	
QY	232	SASAAVIVYNGWSTWENSCASCGRGKRSCTNPAPLNGGAFCEGQNVHDTV 291	
Db	241	STTATVIVYNGWSSWAENSPCNRGCGKRTCTNPAPLNGGAFCEGQ-AFOKTA 299	
QY	292	SSILVSDGWSWPSKWSACGLDCTHWRSCSDPAPRNGGECQGTDLTRNCTSLCV 351	
Db	300	CTTICPDGATEWENKWSACSTCAHWRSCMAPPPQNGGRDCSTGLDLSKNTDGLCM 359	
QY	352	HS-----ASGPEDVALYVGL-IAVAVLVLVLLVLLVYCKKGLSDVADSS-IL 401	
Db	360	QSEFVPAVLEASG--DAALYAGLVAVFVVAAILMAGVVVYRNRCDFTDITDSSAAL 417	
QY	402	TSGQPVSICKSKADNPILL--TTQPDLSITT--TYQSLCPROGSPKFOITGHILSP 458	
Db	418	TGGFHPVNFKARPSPNPQLLPSPPDLTASAGYRGVYALQDS-TDKIPMTNSPLLD 476	
QY	459	LGGRHRLTHSSPT-----SEAEFVSRSLTONY-----FRS----- 490	
Db	477	LPSLVKVIYSSSTGSGFLADGADLLGLVPPGYPSDFARDTHFLHRSASLSGSQLLG 536	
QY	491	LPRGTSNMITYGFNPLGRLMINTGSLILIPDAIPRGIYIEYTLTHKPEDVRLPLA- 549	
Db	537	LPRDPGSSVSGTFGLGRLSIPGTGVSLLVPGAIPQCKFYEMYLINKAEET-LPLSE 595	
QY	550	GCOTLLSIVSCGPPGVLLTPVILLADHCEPSPDSWLSRLKKSCSGSEQDVLHGE 609	
Db	596	GTQTVLSPSVCTGPTGLLLCPVILTMPHCAEVSARDWIFQKTAHQHGEQEVVTLDE 655	
QY	610	EAPSHLYCOLEASAGVFTQLGRFALVGEALSVAARLKLKLLFAPVACTSLEYNIRV 669	
Db	656	ETLNTPCYCLEPRACHILLDQLTGYVFTGESYSRAVKRLQAVFAPALCTSLEYSRV 715	
QY	670	YCLHDTHALKEVVOLEKQGLQIQRVRLHFQDSYNLRLSHIDVPSSLWKSLLVSY 729	
Db	716	YCLEDTVPALKEVLELERTLGGYLVVEEPKPLMFKDSYHNLRLSLHDLPHAWHRSKLLAKY 775	
QY	730	QEIPEYHWTGQVRLHCTFLERSPTSDIACKLWQVGEQGSIFNFINIKTRFF 789	
Db	776	QEIPIYHWSGSQALHCTFLERSHLSASTELTCKICVRQVEGEQIQIQLHTTLA-ETPA 834	
QY	790	AELIALESEAG--VPALVGPFAFKIPFLIRQKIISSLDPPCRRGADWRTLAOKLHSLH 847	
Db	835	GSOLDLCSAPGSTVTTQLGPYAFKIPLSIRQICNSLDAPNSRGNDRMLAKLSMDRYL 894	
QY	848	SFFASKPSPTAMILNWARHPNPNGLQLAARAVAGLQPDAGLFTVSEAE 899	
Db	895	NYFATKASPTGVILDLWEALQDDGLNLSALAEEMGKSEMLVAVATDGD 946	
RESULT 15			
AAW78900			
ID	AAW78900 standard; protein; 943 AA.		
XX			
AC	AAW78900;		

XX	25-MAR-2003 (revised)		
DT	21-DEC-1998 (first entry)		
XX	Rat UNC-5 homologue UNC5H-2.		
DE	UNC5H-2; rat; netrin receptor; cell migration; axon guidance; diagnosis; therapy.		
XX	Rattus sp.		
XX	Key Location/Qualifiers		
FT	Peptide 148..161		
FT	/note= "peptide used to raise rabbit polyclonal antisera"		
FT	Misc-difference 753		
FT	/note= "encoded by CG"		
FT	Peptide 909..924		
FT	/note= "peptide used to raise rabbit polyclonal antisera"		
XX	W09837085-A1.		
XX	27-AUG-1998.		
XX	19-FEB-1998; 98WO-US003143.		
XX	19-FEB-1997; 97US-00808982.		
XX	(REGC) UNIV CALIFORNIA.		
XX	Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Kleinomasu K;		
XX	WPI: 1998-495364/42.		
XX	N-PSDB; AAV52942.		
XX	Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and the biopharmaceutical industry.		
XX	Claim 1; Page 24-26; 32pp; English.		
XX	UNC5H-1 and UNC5H-2 (see AAW78900) are rat homologues of Caenorhabditis elegans UNC-5 protein. Their amino acid sequences were deduced from an isolated unc5H cDNA clones (see AAV52940 and AAV52942) isolated from an E18 brain cDNA library. The predicted proteins show similarity with UNC-5, possess 2 predicted Ig-like domains and 2 predicted thrombospondin type-1 repeats, a predicted membrane spanning region, and a large intracellular domain. They are predicted to be involved in cell migration and axon guidance, and are characterised as receptor proteins for netrin. Human UNC5H-1 (see AAW78899) and UNC5H-2 (see AAW78901) proteins are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly from transfected host cells. The invention also provides unc-5 hybridisation probes and primers, vertebrate UNC-5-specific binding agents such as specific antibodies, and methods of making and using the vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate vertebrate unc-5 gene expression) and in the biopharmaceutical industry (e.g. as immunogens, reagents for modulating cell guidance, reagents for screening chemical libraries for lead pharmacological agents, etc.).		
XX	(Updated on 25-MAR-2003 to correct PI field.)		
XX	Sequence 943 AA;		
XX	Query Match 52.0%; Score 2490; DB 2; Length 943;		
XX	Best Local Similarity 52.3%; Pred. No. 1.1e-206;		
XX	Matches 496; Conservative 144; Mismatches 226; Indels 82; Gaps 19;		
QY	9	PALGIVLAAMLGRSGCAQSAQATVAPVPGANPDLPLPHELVPEPDIYVKNKPVLLVCKAV 68	
Db	21	PSLAGI-----DSGAQ---GLPDSFSPAQBPLFLEPEDAYIVKNKPEVLRCAFF 70	
QY	69	PATQIFFKNGEWSVQVDHVIETSDGSGGETMVRINVSQQVEKVFGLBEYWCQVVA 128	
Db	71	PATQIFYKNGEWSQKHVTKQESLDEATGLRIRVQIEVSRQVVEELFGLDYWCQVVA 130	

this Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 6, 2004, 18:20:44 ; Search time 152 Seconds
(without alignments)
3282.242 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRPLGLWALLGIVLAAML.....AVAGLQPDAGLFTVSEAC 899

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh

-Q=/cgn2_1/USPTO.spool/US09970944/runat_05102004_112011_1655/app.query.fasta_1.1095

-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi

-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptc -NOEM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09970944 @CGN_1_1_75 @runat_05102004_112011_1655 -NCPUS=6 -ICPU=3

-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*

1: /cgn2_6/prodata/2/ina/5A COMB.seq:*

2: /cgn2_6/prodata/2/ina/5B COMB.seq:*

3: /cgn2_6/prodata/2/ina/6A COMB.seq:*

4: /cgn2_6/prodata/2/ina/6B COMB.seq:*

5: /cgn2_6/prodata/2/ina/PCTUS COMB.seq:*

6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4545.5	95.0	3014	2	US-08-808-982-1
2	4545.5	95.0	3014	3	US-09-306-902A-1
3	2694	56.3	1787	2	US-08-808-982-2
4	2694	56.3	1787	3	US-09-306-902A-2
5	2463	51.5	2831	2	US-08-808-982-3
6	2463	51.5	2831	3	US-09-306-902A-3
7	2039	42.6	1282	4	US-09-833-381-1806
8	1123	23.5	1605	4	US-09-833-381-1807
9	600.5	12.5	771	1	US-08-253-155A-17
10	267.5	5.6	5749	4	US-09-262-537-48
11	253	5.3	305	2	US-08-808-982-4
12	253	5.3	305	3	US-09-306-902A-4

13	250	5.2	1326	3	US-08-985-526-4	Sequence 4, Appli
14	250	5.2	7231	4	US-09-919-172-64	Sequence 64, Appli
15	245.5	5.1	657	3	US-08-985-526-2	Sequence 2, Appli
16	211.5	4.4	3874	4	US-09-877-730-31	Sequence 31, Appli
17	204.5	4.3	3991	4	US-08-506-296B-3	Sequence 3, Appli
18	204	4.3	3210	4	US-09-877-730-1	Sequence 1, Appli
19	204	4.3	3453	4	US-09-877-730-7	Sequence 7, Appli
20	192.5	4.0	2715	4	US-09-877-730-5	Sequence 5, Appli
21	192.5	4.0	2958	4	US-09-877-730-9	Sequence 9, Appli
22	186	3.9	1143	4	US-09-877-730-3	Sequence 3, Appli
23	184	3.8	7702	4	US-09-023-655-1336	Sequence 1336, Ap
24	178.5	3.7	3675	4	US-09-930-872-3	Sequence 3, Appli
25	178.5	3.7	4042	4	US-09-930-872-5	Sequence 5, Appli
26	175.5	3.7	8312	4	US-09-620-312D-1048	Sequence 1048, Ap
27	173.5	3.6	6814	4	US-09-484-970B-66	Sequence 66, Appli
28	170.5	3.6	3126	4	US-09-392-184-7	Sequence 7, Appli
29	170.5	3.6	5690	2	US-08-447-464-2	Sequence 2, Appli
30	170.5	3.6	5690	2	US-08-716-679-2	Sequence 2, Appli
31	168	3.5	2114	4	US-09-130-491-7	Sequence 7, Appli
32	168	3.5	4548	4	US-09-571-479C-5	Sequence 5, Appli
33	167.5	3.5	3638	4	US-09-369-364A-8	Sequence 8, Appli
34	166	3.5	4192	4	US-09-122-126B-1	Sequence 1, Appli
35	166	3.5	4192	4	US-09-634-286A-1	Sequence 1, Appli
36	165	3.4	4843	3	US-08-986-485-1	Sequence 1, Appli
37	164.5	3.4	3218	4	US-09-369-364A-6	Sequence 6, Appli
38	163	3.4	4078	4	US-09-016-434-1132	Sequence 1132, Ap
39	163	3.4	6000	1	US-08-348-006B-6	Sequence 6, Appli
40	163	3.4	6000	2	US-08-800-825A-6	Sequence 6, Appli
41	163	3.4	6000	3	US-09-158-657-6	Sequence 6, Appli
42	163	3.4	6000	5	PCT-US94-10166-6	Sequence 6, Appli
43	162.5	3.4	5357	4	US-09-392-184-5	Sequence 5, Appli
44	160.5	3.4	3889	4	US-09-568-559-1	Sequence 1, Appli
45	160.5	3.4	4676	4	US-09-130-491-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-808-982-1

; Sequence 1, Application US/08808982

; Patent No. 5939271

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; APPLICANT: Leonardo, E. David

; APPLICANT: Hink, Lindsay

; APPLICANT: Masu, Masayuki

; APPLICANT: Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION NUMBER:

; APPLICATION NUMBER: US/08/808,982

; FILING DATE:

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: OSMAN, RICHARD A

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: UC96-217

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 343-4341

; TELEFAX: (415) 343-4342

; INFORMATION FOR SEQ ID NO: 1:

```
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 3014 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
US-08-808-982-1
```

```
Alignment Scores:
Pred. No.: 0 Length: 3014
Score: 4545.50 Matches: 852
Percent Similarity: 96.78% Conservative: 19
Best Local Similarity: 94.67% Mismatches: 26
Query Match: 94.96% Indels: 3
DB: Gaps: 3
```

US-09-970-944-2 (1-899) x US-08-808-982-1 (1-3014)

```
QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyValLeuAlaAlaTrpLeu 20
DB 1 ATGGCGCGTCCGGCCCGCGCTGTGGCCAGTGTCTCTGGGCATATGCTCTGGCGCGCTGGCTT 60
QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAlaAsnPro 40
DB 61 CGTGGTTCGGGTGCCCGAGAGTGCACAGGTGGCCAAATCCAGTGGCCGGTGCACACCCC 120
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
DB 121 GACCTGTGCCCGCCACTTCTCTGGTAGAGCTGTAGAGCGTGTACATTTGTCAAGAACAAAGCCG 180
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
DB 181 GTGTGTGTGTGTGCAAGGCTGTGCTCCCGCCAGTCTCTTCAAGTGCATGGGAA 240
QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
DB 241 TGGTTCGCCCGAGTTCGATCATCGTAATTTGAACGAGCAGCAGCAGCAGCGGATTGCCA 300
QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
DB 301 ACCATGAGGTCGTATCAACGTATCGAGCAGCAGGTAGAGAAAGTGTGGGCTGGAG 360
QY 121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
DB 361 GAATACTGTGTGCCAGTGTGGCATGAGTCTCTCGGTTACCAACCAAAATCGAAGGCC 420
QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
DB 421 TACATCCGATTGCCTATTTTGGCAAGAACTTTTGGAGCAGGAGCCACTGGCCAAAGAGTG 480
QY 161 SerLeuGlnGlnGlyIleValLeuProCysArgProGluGlyIleProProAlaGlu 180
DB 481 TCATGAGCAAGCAATGTACTACCTTGTGCGCCCGCCAGAGAAATCCCGCCAGCTGAG 540
QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
DB 541 GTGGAGTGGCTTCGAATAGAGACTCTGTGGACCCCTCTCGATCCCAATGTGTATC 600
QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
DB 601 ACGCGGAGCAGCAGCCCTAGTCTGTGCGTCAAGCGCGCTGGCCGACACGGCCAACTACAC 660
QY 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240
DB 661 TGTGTGCCCAAGAACATCGTAGCCGCTGCGGAAGACCTCTGAGCGGGTCAATGTATTAT 720
QY 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
DB 721 GTGNACCGTGGTGGTGCAGCTGAGTGTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGCGT 780
QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyAlaPhe 280
DB 781 GGCTGGCAGAAACGAGCGGAGTGCACCAACCGCGCACCTCTCAACGGGGCGCGCTTC 840
```

```
QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
DB 841 TGTGAGGGGAGATGTC---CAGAAAACAGCGCTGGCCACTCTGTGCCAGTGGATGGG 897
QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
DB 898 AGCTGGAGTTCGTGGAGTAAGTGGTTCAGCTGTGGGCTTGAAGTGCACCCACTGGCGGAGC 957
QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
DB 958 CGCGAGTGTCTGTACCCAGCAGCCCGCAATGGAGGTGAGAGTGTGGGGTGTGACCTG 1017
QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
DB 1018 GACACCCGCACTGTACCAAGTACCTTCGCTGCACACCGCTTCTTGGCCGAGGAGCTG 1077
QY 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
DB 1078 GCTCTCTACATCGCCCTTGTGCTGTGGTGTGTGCTCTTCTTGTGTGTGGCCCTT 1137
QY 381 IleLeuValTyrCysArgLysLysGlyLeuAspSerAspValAlaAlaAspSerIle 400
DB 1138 GGACTCATTTACTGTCTGCAAGAGGAGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197
QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAlaAspAsnProHisLeu 420
DB 1198 CTCACCTCGGGCTTCAGCTGTACGATCAAGCCAGCAAGCAGCAGCAACCCACCTG 1257
QY 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrGlnGlySerLeuCysPro 439
DB 1258 CTCACCATCCAGCCAGCAGCTCAGCACCACCTACCACTTACCAGGCGAGTCTATGTTCG 1317
QY 440 ArgGluAspGlyProSerProLysPheGlnLeuThrAsnGlyHisIleLeuSerProLeu 459
DB 1318 AGCAGGAGTGGACCCAGCCCAAGTTCAGCTCTTAATGGTCACTGTCTCAGCCCACTG 1377
QY 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
DB 1378 GGGAGTGGCGGCATAGTTGCACACAGCTCACCACCTCTGAGGCTGAGGACTTGTGC 1437
QY 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
DB 1438 TCCCGCTCTCCACCCCAAACTACTTTCGTTTCCTGCTGCGCGCGCAGCAGCAACATGCC 1497
QY 500 TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
DB 1498 TACGGGACCTTCAACTTCTCGGGGCGCGCTGATGATCCCTAATACGGGATCAGCCTC 1557
QY 520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
DB 1558 CTATACCCCGGATGCCATCCCGGAGGAGAAAGATCTACGAGATCTACCTCACATGCAC 1617
QY 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
DB 1618 AAGCCAGAACAGCTGAGTGTGCCCTAGCTGGCTGTGTCAGACCCCTGTGAGTCCAGCTGT 1677
QY 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
DB 1678 AGCTGTGGGCGCCAGAGTCTGCTCACCAGGCGAGTCACTCTGATATGGACCTGT 1737
QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
DB 1738 GGAGAGCCAGCCCTGACAGTGGAGTCTGGCGCTCAAAAAGCAGTCTCTCGAGGGCAGT 1797
QY 600 TrpGluGlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGln 619
DB 1798 TGGGAG---GATGTGTGTGACCTTGTGTGAGGAGTCACTTCCCACTCTACTACTGCCAG 1854
QY 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
DB 1855 CTGAGGCGCGGGCGCTGTATGTCTTCCAGGAGCAGCTGGCGCGCTTTCCTGTGTAGGA 1914
```



```
QY 640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla 659
Db 1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCTCAGGCTCTCTTGTCTTCCGTGGCC 1974
QY 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db 1975 TGTACGTCTCTTGTAGTACACATCCAGTGTACTGCTTACAGACACCCACGAGCTCTC 2034
QY 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnLeuProArgVal 699
Db 2035 AAGGAGTGGTGCAGCTGGAGAAGAGCTAGGTGGACAGCTATCCAGGAGCTCGCGTC 2094
QY 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db 2095 CTGCACCTTCAAAGACAGTTTACCACACCTTACGTCCTCCATCCAGCGTCCAGCTCC 2154
QY 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db 2155 CTGTGGAGAGCAAGCTACTTGTCTGACTACAGGAGATCCCTTTTACCACATCTGGAAC 2214
QY 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db 2215 GGCACCCAGCATCTGCACTGCACCTTCACTTCCCTGGAGCGCATCAACGCCAGCACCAGC 2274
QY 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db 2275 GACCTGGCTGCAAGGTGTGGTGTGGCAGGTGGAGGAGATGGGAGAGCTTCAACATC 2334
QY 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db 2335 AACTTCAACATCACTAAGGACACAAGGTTTGTGTAATTGTGGCTCTGGAGAGTGAAGG 2394
QY 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db 2395 GGGGTCCAGCCAGCTGTGGGCCCGCCAGTGCCTTCAAGATCCCTTCTATTCCGCCAAG 2454
QY 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db 2455 ATCATCGCCAGTCTGGACCCACCTTGCAGCCGGGGCGGAGTGAAGAAGTCTAGCCGAG 2514
QY 840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db 2515 AAACCTTCACTGGACAGCATCTTAGCTTCTTTGGCTTCAAGCCAGCCGCTACAGCCATG 2574
QY 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaIa 879
Db 2575 ATCTTCACTATGGAGGACGCGGACCTTCCCAACGGCAACTTCCGCGAGCTGGCAGCA 2634
QY 880 AlaValAlaGlyLeuGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 2635 GCTGTGGCCGAGTGGGCCAACACAGATGCTGGCTTCTTACGGTCTGGAGGCCGAGTGT 2694
```

RESULT 2

```
US-09-306-902A-1
; Sequence 1, Application US/09306902A
; Patent No. 6277585
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/306,902A
FILING DATE: 07-May-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC96-217
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3014 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-306-902A-1

Alignment Scores:
Pred. No.: 0 Length: 3014
Score: 4545.50 Matches: 852
Percent Similarity: 96.78% Conservative: 19
Best Local Similarity: 94.67% Mismatches: 26
Query Match: 94.96% Indels: 3
DB: 3 Gaps: 3

US-09-970-944-2 (1-899) x US-09-306-902A-1 (1-3014)
QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db 1 ATGGCCGTCCGGCCCGGCTGTGGCCAGTGTCTTGGGCATAGTCTCTGCCGCCCTGGCTT 60
QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db 61 CGTGTTCGGGTGCGCAGCAGAGTGCACGGTGGCCAATCCAGTCCCGGTCGCAACCC 120
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db 121 GACCTGTGCCCATCTTCTGTAGAGCTGAGGACGTGTACATGTCAAGAACAAAGCGG 180
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db 181 GTGTGTGGTGTGCAAGGCTGTGCTGCCACCCAGATCTTCTTCAAGTGCATGGGNA 240
QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db 241 TGGGTCCGCCAGTTCGATCAGCTAATTGAACGACGACGACGACGAGGATGGCA 300
QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db 301 ACCATGGAGGTCGGTATCAAGTATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db 361 GAATACTGGTCCAGTGTGTGGCATGGAGCTCTCTCGGATACCAACCAAAAGTCAGNAGCC 420
QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db 421 TACATCCGATTGGCTATTTCGCAAGAACCTTTGAGCAGGAGGAGGAGGAGGAGGAG 480
QY 161 SerLeuGlnGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db 481 TCACTGGAGCAAGCATGTACTACCTTGTGCCGCCGCCAGAGGAATCCCCCGAGCTGAG 540
QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db 541 GTGGAGTGGCTTCGAATAGAGACCTGTGGACCCCTCCCTCGATCCCATGTGTATCATC
```

QY	201	ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr	220	QY	560	SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys	579
Db	601	ACGGGGAGCACAGCCTAGCTCGCTCAGGCCGCTGGCCGACACGGCCAACTACACC	660	Db	1678	AGCTGTGGCCCCCAGGAGTCTGTCTACCCGCCAGTCATCTTGCATGAGCACCTGT	1737
QY	221	CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr	240	QY	580	GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer	599
Db	661	TGTGTGGCCCAAGAAATCATGAGCCGTCGCCGAGCACCTCTGCAGCGGTCAATGTTAT	720	Db	1738	GGAGAGCCAGCCCTGACAGCTGGAGTCTGGCCCTCAAAAAGCAGTCTCTCGAGGCACT	1797
QY	241	ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg	260	QY	600	TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln	619
Db	721	GTGAAGGTGGTGGTTCGAGCGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	780	Db	1798	TGGGAG--GATGTGTGTGACACCTTGTGAGGAGTCACTTCCACCTCTACTACTGCCAG	1854
QY	261	GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe	280	QY	620	LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGluArgPheAlaLeuValGly	639
Db	781	GGCTGGCAGAAACGGAGCGGAGCTGCACCAACCGGCACCTCTCAACGGGGCGCCTTC	840	Db	1855	CTGGAGCGCGGGCCTGTATGTCTTCACGAGAGCTGGCGCGCTTTCCTTGGTAGGA	1914
QY	281	CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly	300	QY	640	GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla	659
Db	841	TGTGAGGGCAGAAATGTC--CAGAAAACAGCTGGCCCACTCTGTGCCAGTGGATGG	897	Db	1915	GAGGCCCTCAGCGTGGCTGCCACCAAGCGCTCAGGCTCTTCTGTCTCCCGTGGCC	1974
QY	301	SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer	320	QY	660	CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu	679
Db	898	AGCTGGAGTTCTGGAGTAAAGTGTTCAGCTGTGGGCTTGACTGCACCCACTGGCGAGC	957	Db	1975	TGTACGTCCTTTCAGTACAAATCCGAGTGTACTGCCTACACGACACCCACGCGCTTC	2034
QY	321	ArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGlnGlyThrAspLeu	340	QY	680	LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal	699
Db	958	CGCGAGTGTCTGACCCAGCACCCCGCAATGGAGTGGAGGTTCTGGGGTGTGACCTG	1017	Db	2035	AAGGAGTGGTGCAGCTGGAGAGCAGTAGTGTGACAGCTGATCCAGAGGCTCGCGTC	2094
QY	341	AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal	360	QY	700	LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer	719
Db	1018	GACACCGCACTGTACAGTAGTACCTCTGCTGTCACACCGCTTCTTGGCCCGAGGACGTG	1077	Db	2095	CTGCACCTCAAAGACAGTTTACCACAACTTACCTCTCTCCATCCACGCTGCCAGCTCC	2154
QY	361	AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu	380	QY	720	LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn	739
Db	1078	GCTCTCTACATGGCCTTGTGCTGTGGTGTGTGCTCTTCTTGTGTGTGTGGCCCTT	1137	Db	2155	CTGTGGAGAGAGAGTACTTGTCTAGCTACAGGAGATCCCTTTTACCACACTCGAAC	2214
QY	381	IleLeuValTyrCysArgLysLysGlyLeuAspSerAspValAlaAspSerSerIle	400	QY	740	GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer	759
Db	1138	GGACTCATTTACTGTCCGAGAGGAGGAGGCTGGACTCCGATGTGGCCGACTGCTCCATC	1197	Db	2215	GGCACCAGCAGTATCTGACCTGCACCTTACCTTGGAGCGCATCAACGCCAGCACCGC	2274
QY	401	LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu	420	QY	760	AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle	779
Db	1198	CTCACCTCGGGCTTCAGGCTGTACAGCATCAAGCCAGCCAGCAACACACCCACCTG	1257	Db	2275	GACCTGGCTCGAAGGTGGGTGGCAGGTGGAGGAGATGGGCAGAGCTTCAACATC	2334
QY	421	LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro	439	QY	780	AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla	799
Db	1258	CTCACCATCCAGCCAGCTCAGCACCACTTACCACCTACCAGGCGAGTCTATGTTCTG	1317	Db	2335	AACTTCAAACATCACTAAGGACACAAGGTTTGTGTAATTTGTGGCTCTGGAGAGTGAAGG	2394
QY	440	ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisIleLeuSerProLeu	459	QY	800	GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys	819
Db	1318	AGGAGAGTGGACCCAGCCCAAGTTCCAGCTCTCTAAATGGTCACTGTCTCAGCCCACTG	1377	Db	2395	GGGGTCCCAAGCCTGGTGGGCCAGTGTCTTCAAGATCCCTTCTCATTCGGCAAAAG	2454
QY	460	GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluPheVal	479	QY	820	IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln	839
Db	1378	GGGAGTGGCGGCATACGTTGACCAACAGCTACCCACCTCTGAGGCTGAGGACTTCGTC	1437	Db	2455	ATCATCGCCAGTCTGGACCCACCTGTCAGCGGGCGCGACTGGAGAACTCTAGGCCAG	2514
QY	480	SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr	499	QY	840	LysLeuHisIleAspSerHisIleuSerPhePheAlaSerLysProSerProThrAlaMet	859
Db	1438	TCCCGCTCTCCACCCAAAACACTACTTCTGTTCTGCTCCCGCGCCACCAACATGGGC	1497	Db	2515	AAACTTCACCTGACAGCCATCTTACTTCTTGTGCTTCAAGCCCGCCCTCAGGCCATG	2574
QY	500	TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu	519	QY	860	IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla	879
Db	1498	TAGGGACCTTCAACTTCTTCGGGGCGCGCTGATGATCCCTAATACGGGATCAGCCCT	1557	Db	2575	ATCCTCAACTATGGAGGCGACGGCACTTCCCAACGGCACTCGGCCAGCTGGCAGCA	2634
QY	520	LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis	539	QY	880	AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys	899
Db	1558	CTCATACCCCGGATGTCATCCCGCGAGGAAAGATCTACGAGATCTTACCTCACACTGCAC	1617	Db	2635	GCTGTGGCGGAGTGGCCCAACAGATGTGGCTCTTCAACGCTGTGGAGGCGAGTGT	2694
QY	540	LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal	559				
Db	1618	AAGCCAGAGACGTGAGTTGCCCTAGCTGGCTGTGACAGCCCTGTGTGAGTCCAGTCGTT	1677				

RESULT 3

US-08-808-982-2

; Sequence 2, Application US/0808982

; Patent No. 5939271

; GENERAL INFORMATION:

APPLICANT: Tessier-Lavigne, Marc
 APPLICANT: Leonardo, E. David
 APPLICANT: Hink, Lindsay
 APPLICANT: Masu, Masayuki
 APPLICANT: Kazuko, Keino-Masu
 TITLE OF INVENTION: Netrin Receptors
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 STREET: 268 BUSH STREET, SUITE 3200
 CITY: SAN FRANCISCO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/808,982
 FILING DATE:
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: OSMAN, RICHARD A
 REGISTRATION NUMBER: 36,627
 REFERENCE/DOCKET NUMBER: UC96-217
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 343-4341
 TELEFAX: (415) 343-4342
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1787 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-808-982-2

Alignment Scores:

Pred. No.: 8,576-252 Length: 1787
 Score: 2694.00 Matches: 543
 Percent Similarity: 97.14% Conservative: 1
 Best Local Similarity: 96.96% Mismatches: 11
 Query Match: 56.28% Indels: 11
 DB: 2 Gaps: 2

US-09-970-944-2 (1-899) x US-08-808-982-2 (1-1787)

344 AsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363
 3 AACGTACAGTACCTCTG-GTACACAGTCTTCTGCCCTGAGGAGCTGGCCCTCTAT 61
 364 ValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeuVal 383
 62 GTGGGCTCATCGCGTGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
 384 TyrCysArgLysLysGlyGlyLeuAspSerValAlaAspSerSerIleLeuThrSer 403
 122 TATTGCCGGAAGAGAGGGGCTGACTCAGATGTGGCTGACTCGTCCATTCTACCTCA 181
 404 GlyPheGlnProValSerIle-LysProSerIysAlaAspAsnProHisLeuThrIle 423
 192 GGTTCAGCCCGCTCAGCATCTAAGCCAGCAAGACACACCCCTGCTGCTGCTGCTGCT 241
 423 eGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThrThrThrThr 442
 242 CCAGCCGACCTCAGCACCAACACACCTACAGAGGAGCTCTCTGCTCCCGGAGGA 301
 442 pGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGly 462
 302 TGGGCCAGGCCCAAGTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGGG 361

462 YArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSerArgLe 482
 362 CGCCACACACTGCACACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCGCGCT 421
 482 uSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetTyrTyrGlyTh 502
 422 CTCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACAGCAATGACCTATGGGAC 481
 502 rPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuPr 522
 482 CTTCAACTTCCTCGGGGCGCGCTGATGATCCCTAATACAGGAATCAGCCTCTCATCCC 541
 522 cProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGl 542
 542 CCCAGATGCCATACCCCGAGGAGAGATCTATGAGATCTACCTCAGCTGCACAGCCCGGA 601
 542 uAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGl 562
 602 AGAGCTGAGGTTGCCCTAGCTGGCTGTGAGCCCTGCTGAGTCCCATCGTAGCTGTGG 661
 662 ACCCTCGCGCTCTGCTCAGCCGCGCAGTATCTGCTATGGACCACTGTGGGGAGGC 721
 582 yProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluPr 582
 662 oSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGl 602
 722 CAGCCCTCAGCAGCTGGAGCCTGGCCCTCAAAAGCAGTCGTGCGAGGAGGC-TGGAG-- 778
 602 nAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAl 622
 779 -GATGT-CTGCACCTGGCGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGC 836
 622 aserAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLe 642
 837 CAGTGCCTGCTACGTCCTTCAACGAGCAGCTGGCGCGCTTTCCTCTGGTGGGAGAGCCCT 896
 642 uSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAlaCysThrSe 662
 897 CAGGTGGCTCGCGCAAGCGCTCAAGCTGCTTCTGTGTGGCGCGGTGGCGCTGCACCTC 956
 662 rLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluVa 682
 957 CCTCGAGTACACATCCGGGTCTACTGCTGATGACACCCACCATGATGACACTCAAGAGAGT 1016
 682 lValGlnLeuGlyLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPh 702
 1017 GGTGCAGCTGGAGAAAGCAGCTGGGGGAGCAGCTGATCCAGGAGCCAGCGGTCTGCACTT 1076
 702 eLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLys 722
 1077 -AAGGACAGTTACCACACCTGCC-CTATC-ATCCACGATGTGCCAGCTCCCTGTGGAA 1133
 722 sSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTyrAsnGlyThrGl 742
 1134 GAGTAAGCTCTTGTGAGTACAGGAGATCCCTTTTATCACATCTGGAATGGCACCGCA 1193
 742 nArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAl 762
 1194 CGCGTACTTGCACCTGCACCTTCACCTGGAGCGTGTGAGCCCGCAGCACTAGTAGCTGGC 1253
 762 aCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAs 782
 1254 CTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGAGCGGAGAGCTTCAGCATCAACTTCAA 1313
 782 nIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValPr 802
 1314 CATCACCAAGACACAGGTTGTGTGAGTGTGCTGCTGAGAGTGAAGGGGGGTGCC 1373
 802 oAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleLeSe 822
 1374 AGCCCTGTGGGCCCGAGTGCCTTCAAGATCCCTTCTCTATTCCGAGAGATTAATTTC 1433
 822 rSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHi 842

QY	722	sSerLysLeuLeuValSerTyrGlnGluIleProPhePyrHisIleTrpAsnGlyThrGl	742
Db	1134	GAGTAAGCTCTCTGTCACTACCAAGAGATCCCTTTTATCACATCTGGAATGGCACGCA	1193
QY	742	nAtqTyrLeuHisCystrPheThrLeuGluAtrqValSerProSerThrSerAspLeuAl	762
Db	1194	CGGTACTTGCACATGCACCTTCACCTCGAGCGTGTGAGCCCAAGCACTAGTGAACCTGGC	1253
QY	762	aCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAs	782
Db	1254	CTGCAAGCTCTGGGTGTGGCAGGTGGAGGGCAGCGGACAGCTTCACGATCAACTTCAA	1313
QY	782	nIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValPr	802
Db	1314	CATCACC AAGGACACAAGGTTTGTGAGCTGTGGCTCTGGAGAGTGAAGCGGGGTCCC	1373
QY	802	oAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnIlyIleIleSe	822
Db	1374	AGCCCTCGTGGGCCCAAGTGCCTTCAAAGATCCCTTCCTCATTCGGCAGAGATAATTC	1433
QY	822	rSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrIleuAlaGlnLysLeuHi	842
Db	1434	CAGCTCGACCCACCCCTGTAGCGGGGTGGCACTGGCGGACTGTGCCCAGAAACTCCA	1493
QY	842	sLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAs	862
Db	1494	CCTGGACAGCCATCTCAGCTTCTTTGGCTCCAGGCCAGGCCACAGCCATGATCTTCAA	1553
QY	862	nLeuTrpGluAlaAargHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAl	882
Db	1554	CTTGTGGAGGCGGGACATCCCCAACGCAACTCAGCCAGCTGGCTGCAGCAGTGGC	1613
QY	882	aGlyLeuGlyGlnProAspAlaGlyLeu-PheThrVal-SerGluAlaGluCys	899
Db	1614	TGGGATGCGCCACAGACCGTGGCTCTTTTCAGTGTTCGAGGCTGAGTGC	1667

RESULT 5

US-08-808-982-3
; Sequence 3, Application US/08808962
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:


```

;
; TITLE OF INVENTION: Netrin Receptors
;
; NUMBER OF SEQUENCES: 9
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
;
; ZIP: 94104
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: Patent In Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/09/306,902A
;
; FILING DATE: 07-May-1999
;
; CLASSIFICATION: <Unknown>
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME: OSMAN, RICHARD A
;
; REGISTRATION NUMBER: 36,627
;
; REFERENCE/DOCKET NUMBER: UC96-217
;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: (415) 343-4341
;
; TELEFAX: (415) 343-4342
;
; INFORMATION FOR SEQ ID NO: 3:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 2831 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: double
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: cDNA
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
;
US-09-306-902A-3

```

Alignment Scores:

```

Pred. No.:      5,72e-229      Length:      2831
Score:          2463.00      Matches:      495
Percent Similarity: 67.41%      Conservative: 144
Best Local Similarity: 52.22%      Mismatches: 227
Query Match:      51.45%      Indels:      83
DB:              3          Gaps:      19

US-09-970-944-2 (1-899) x US-09-306-902A-3 (1-2831)

QY 9 ProAlaLeuLeuGluGlyIleValLeuAlaLaTrpLeuArgGlySerGlyAlaGlnSer 28
D 61 CCGAGCTTAGCAGGCAATT-----GACTCTGGTGCCAG----- 93
QY 29 AlaThrValAlaAsnProValProGlyAlaAsnProAspLeuProHisPheLeuVal 48
D 94 ---GGACTCCCGAGACTCTCCATCAGCACC CGCGGAGCAGCTGCCTCATTCTCTGCTG 150
QY 49 GluProGluAspValTyrIleValLysAsnLysProValLeuLeuValCysLysAlaVal 68
D 151 GAACCCAGAGGATGCCTACATCGTAAAGAACAGCCAGTGGAAATTCACCTGCCGAGCCTTC 210
QY 69 ProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArgGlnValAspHisVal 88
D 211 CCTGCCACACAGATCTACTTCAAGTGTAATGCGAGTGGGTAGCCAGAAAGGCCACCGTC 270
QY 89 IleGluArgSerThrAspGlySerSerGlyGluProThrMetGluValArgIleAsnVal 108
D 271 AGCGAGGAGGCTCGATGAGCCCGAGCTTGGAAATACGAGAGGTGCGAGAGGTG 330
QY 109 SerArgGlnGlnValGlnLysValPheGlyLeuGluGluTyrTrpCysGlnCysValAla 128
D 331 TCGCGGCAGCAGGTGGAGAACTTTTGGCTCGAGGACTACTGGTGTCTCAGTGGCTGGCC 390
QY 129 TrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAlaArgLeuArg 148
D 391 TGGAGCTCTTCGGGAACACCAAGAGTCCCGAGCTACATCCGCAATTGCCTACTTGGCC 450

```

```

QY 149 LysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGluGlnGlyIleValLeu 168
D 451 AAGAACTTTGACGAGGACCTCTGGCGAAGAGGTTACCTTGGATCATGAGTCTCTTCTG 510
QY 169 ProCysArgProProGluGlyIleProProAlaGluValGluTrpLeuArgAsnGluAsp 188
D 511 CAGTGGCCGCCACACAGAGGAGTGCTCTGTGGCTGAGGTGGAATGCTCAAGAAATCAAGAT 570
QY 189 LeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSerLeuValVal 208
D 571 GTCATCGATCCCGCTCAGGACACTAACTTCTCTGCTCACTTGCACCAACATCATCATC 630
QY 209 ArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsnIleValAla 228
D 631 CGCAGAGCGCGCTCTCAGACACAGCCAACTACACTGTGTGGCAAAGAATATTGTGGCC 690
QY 229 ArgArgSerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrpSerThrTrp 248
D 691 AAGCGCGGAGCAGCAGCGCCACAGTCATCTCTATGTGAACGGAGGTTGGTCCAGCTGG 750
QY 249 ThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSer 268
D 751 GCAGAAATGGTCACCTCTCTAAACCGCTGCGCGCGAGGTTGCGACAAACGTAAGGACC 810
QY 269 CysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsnValHisAsp 288
D 811 TGCACCAACCCAGCCCCACTCAATGGAGGTGCTCTTCTCGAGGGACAG---GCTTGCAC 867
QY 289 ArgThrValSerSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrp 308
D 868 AAGACGGCTTGCCACCCGCTGTCGCCAGTGGATGGAGCGTGGACTGAGTGGAGCAAGTGG 927
QY 309 SerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCysSerAspProAlaPro 328
D 928 TCCGCTGCAGCAGCAGAGTGTGCGCACTGCGCGCAGCCGCGAGTGTGATGCGCCCGGCC 987
QY 329 ArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAsp 348
D 988 CAGAACGGAGGCGGTGACTGACGCGGAGCGCTACTTGTGATCCAGAACTGACCCGATGGG 1047
QY 349 LeuCysVal-----HisSerAlaSerGlyProGlu----- 358
D 1048 CTGTGCGTGCTGAATCAGAGAACTCTAAACGACCCCTAAAGCGCGCCCTGGAGCGCTCG 1107
QY 359 ---AspValAlaLeuTyrValGlyLeu---IleAlaValAlaValCysLeuValLeuLeu 376
D 1108 GGAGACGTGGGCGCTGTATCGGGCGCTCGTGTGGCGCGCTCTTTGGTCTTCGACATTCTC 1167
QY 377 LeuLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAla 396
D 1168 ATGGCTGTAGAGTGATCGTGTACCGGAGAAACTGCCGGGACTTCGACGAGACATCACT 1227
QY 397 AspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSerLysAla 415
D 1228 GACTCTCTGTGCGCTCACTGTGTTCACCCCGCTCAACTTCAAGACTGCAAGGCC 1287
QY 416 AspAsnProHisLeuLeu-----ThrIleGlnProAspLeuSerThrThr----- 432
D 1288 AGCAACCCACAGCTCTCGACCCATCCGCGCTCCGACCTAAACGGCGGAGTGTGGCATC 1347
QY 433 TyrGlnGlySerLeuCysProArgAlaAspGlyProSerProLysPheGlnLeuThrAsn 452
D 1348 TACCGCGAGCTGTGTATGCGCTCGAGGACTCT---GCCGACAGATCCCTATGACTAAT 1404
QY 453 GlyHisLeuLeuSerProLeu----- 459
D 1405 TCACCCCTTCTGGATCCCTTGCACGACCTCAAGATCAAGTCTATGATCCAGACCATC 1464
QY 460 -----GlyGlyGly----- 462
D 1465 GGCTCTGGGGCTGGCTGGCTGATGAGCGACCGACCTGCTGGGTGTCTTACCACCCGGTACA 1524
QY 463 -----ArgHisThrLeuHisSerSerProThrSerGlu 474

```


Db 1525 TACCAGGGGATTCTCTCCGGGACACCCACTTCTCGCACTCGCGAGC----- 1572
QY 475 AlaGluGluPheValSerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGly 494
Db 1573 -----GCCAGCCCTGGTTCACAG---CACCTCTCGGCGCTCCCTCGAGAC 1614
QY 495 ThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetLeuProAsn 514
Db 1615 CCAGCAGCAGGTGTCAGTGGCACTTGTGTGCTGGGTGGAGGTGACCATTCCTCCGGC 1674
QY 515 ThrGlyIleSerLeuLeuLeuProProAspAlaIleProArgGlyLysIleTyrGluIle 534
Db 1675 ACAGGGGTGAGCTGTGTGATCCAAATGAGCCTATCCCGGGCAAGTCTATGACTG 1734
QY 535 TyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla---GlyCysGlnThr 553
Db 1735 TATCTACGTATCAACAAGACTGAAAGCACC---CTCCCACTTTGGGAAGTTCCAGACA 1791
QY 554 LeuLeuSerProIleValSerCysGlyProProGlyValLeuLeuThrArgProValIle 573
Db 1792 GTATTGAGCCCTCGGTGACCTGCGGGCCACGGGCTCTCTGTGCGCGCTGTGTTC 1851
QY 574 LeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLys 593
Db 1852 CTCACGTGCCCCACTGTCTGAGTCATTGCGGAGACTGGATCTTCCAGCTCAAGACC 1911
QY 594 GlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGluAlaProSer 613
Db 1912 CAGGGCCATCAGGGCCACTGGGAG---GAGGTGCTGACTTTGGATGAGGAGACTCTGAAC 1968
QY 614 HisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGly 633
Db 1969 ACCCTCTGCTACTGCGAGTAGAGCTAAATCTGCCACATCTGTGTGGACCACTGGGT 2028
QY 634 ArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeu 653
Db 2029 ACCTAGCTGTTCAGGGCGAGTCTTACTCCCGCTCCGAGTCAAGCGGTCCAGTAGCC 2088
QY 654 LeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHis 673
Db 2089 ATCTTGGCCCGACCCCTCTGCACCTCCCTGGAGTATAGTCTCAGGGTCTACTGTGGAG 2148
QY 674 AspThrHisAspAlaLeuLysGluValValGlnLeuGlyLysGlnLeuGlyGlnLeu 693
Db 2149 GACACTCTCGAGCACTGAAGAGGTCTTAGAGCTGGAGGAGACTTGGGTGGCTACTTG 2208
QY 694 IleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIle 713
Db 2209 GTGGAGGAGCCCAAGACTTTGCTCTTTAAGGACAGTTACCAACACTACGCTC-TCCTC 2267
QY 714 HisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyrGlnGluIlePro 733
Db 2268 CATGACATCCCCCATGCCCACTGGAGGACAAACTACTGGCAAGTACCAGGAGATTCCC 2327
QY 734 PheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArg 753
Db 2328 TTCTACCATGTGGNACGGGAGCCAGAAAGCCCTGCAGTGCATTTTACCTTGAGAGA 2387
QY 754 ValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAsp 773
Db 2388 CATAGCTAGCTCCACTGAGTTACCTGTAAGGTCTGGTCCGGCAGGTAGAGGGAA 2447
QY 774 GlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeu 793
Db 2448 GCCCAGATTTTCAGTGCACACCAAGCTGGGT---GAGACCCCTGCTGGCTCCCTGGAT 2504
QY 794 AlaLeuGluSerGluAlaGlyValProAlaLeu-----ValGlyProSerAlaPheLys 811
Db 2505 GCACCTGTCTGCCCCCTGGCAATGCTGCCACACACAGCTGGAGCCCTATGCCCTCAAG 2564
QY 812 IleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProProCysArgArgGly 831

Db 2565 ATACCACGTCTCCATCCGCGACAGAGATCTGCAACAGCCTGGACGCCCACTCACGGGGC 2624
QY 832 AlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAla 851
Db 2625 AATGACTGGCGGTCTGTGGCAGACAGAGCTCTCCATGACCGGTACCTGAACACTACTTCGCC 2684
QY 852 SerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsn 871
Db 2685 ACCAAGCTAGTCCACAGGCGTGATCTTAGACCTCTGGGAAGCTCGGCAGCAGGATGAT 2744
QY 872 GlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeu 891
Db 2745 GGGGACCTCAACAGCCTGGCCAGTGCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGCTA 2804
QY 892 PheThrValSerGluAlaGluCys 899
Db 2805 GCCATGACCCTGATGGCGATTGC 2828
RESULT 7
US-09-833-381-1806
; Sequence 1806, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1806
; LENGTH: 1282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1282)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1806
Alignment Scores:
Pred. No.: 2,49e-188 Length: 1282
Score: 2039.00 Matches: 410
Percent Similarity: 95.58% Conservative: 1
Best Local Similarity: 95.35% Mismatches: 11
Query Match: 42.59% Indels: 9
DB: 4 Gaps: 2
US-09-970-944-2 (1-899) x US-09-833-381-1806 (1-1282)
QY 476 GluGluPheValSerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThr 495
Db 3 GAGGAGTTTCGTCTCCCGCTCTCCACCAGAACTACTTCCGCTCCCTGCCCGAGGCACC 62
QY 496 SerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThr 515
Db 63 AGCAACATGACCTATGGGACCTTCAACTTCTCGGGGCGCGCTGATGATCCCTAATACA 122
QY 516 GlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyr 535
Db 123 GGAAATGACCTCTCATCCCCCAGATGCCATACCCCGAGGAGAGATCTATGAGATCTAC 182
QY 536 LeuThrLeuHisLysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeu 555
Db 183 CTCAGCTGCAACAGCCGAAAGCTGAGGTTCCTAGCTGGCTGTGAGACCTCTGCTG 242
QY 556 SerProIleValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeu-A 575
Db 243 AGTCCCATCTGTAGTGTGGACCCCTGGGCGTCTGTCTACCCGCCAGTCACTCTGGG 302
QY 575 laMetAspHisCysGlyGluProSerProAspSerTrp-SerLeuArgLeuLysLysGln 594

```
|||||
Db 303 GTATGGACCATGTGGGAGCCAGCCCTGACAGCTGGAGGCTGGCTCAAAAGCAG 362
Qy 595 SerCysGluGlySerTrpGluGlnAspVal-LeuHisLeuGlyGluGluAlaProSerHi 614
Db 363 TCGTCGAGGGGAGCTGGGAG--GATGTGCTTGACCTGGGCGAGGAGGCCCTCCCA 419
Qy 614 sLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyAr 634
Db 420 CCTCTACTACTGCCAGCTGGAGCCAGTGCCTGCTACGCTTTCACCGAGCAGCTGAGCGG 479
Qy 634 gPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaAlaAlaAlaAlaAlaAla 654
Db 480 CTATGCCCTGTGGGAGGAGCCCTCAGCGTGCTGCCGCAAGCGCCTCAAGCTGCTTCT 539
Qy 654 uPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAs 674
Db 540 GTTTGCGCGGTGGCTGCACCTCGCTGAGTACACATACATCTGCTCTACTGCTGCATGA 599
Qy 674 pThrHisAlaLeuLysGluValValClnLeuGlnLysGlnLeuGlyGlyGlnLeuIl 694
Db 600 CACTCACCAGTGCACCTCAACGTAGTGGTGAGCTGGAGAGCAGCTGGGGGGACAGCTGAT 659
Qy 694 eGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHi 714
Db 660 CCAGGAGCCAGGGTCTGCACTTCACGAGCAGTATCCACACCTGCGCTATCCATCA 719
Qy 714 sAspValProSerSerLeuTrpLysSerLysLeuValSerTyrGlnGluIleProPh 734
Db 720 CGATGTGCCAGCTCCCTGTGGAAGTAAAGTCTCTGTGTCAGTACCAGGAGATCCCTT 779
Qy 734 eTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgVa 754
Db 780 TTATCACATCTGGAAATGGACGACGCGTACTTGCACTGCACCTTCACCTGGAGCGTGT 839
Qy 754 lSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGl 774
Db 840 CAGCCCCAGCACTAGTACCTGGCTGCAAGCTGTGGTGAGGAGTGAGGGGCGAGCG 899
Qy 774 yGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAl 794
Db 900 GCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAAGGTTTGCTGAGCTGTGC 959
Qy 794 aLeuGluSerGluAlaGlyValProAlaLeuValGlyProSerAlaPheLysIleProPh 814
Db 960 TCTGAGAGTGAAGCGNGGTCCTCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTT 1019
Qy 814 eLeuIleArgGlnLysIlelle-SerSerLeuAspProProCysArgArgGlyAlaAspT 834
Db 1020 CCTCATTCGGCAGAGATAATTTTCCAGCTGGACCCACCTGTAGGGGGGTGCGACT 1079
Qy 834 rpArg-ThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLys 853
Db 1080 GCGGAACCTCTGGCCAGAAATCCACCTGGACAGCCATCTCAGCTTCTTTGCCCTCCAAG 1139
Qy 854 ProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsn 873
Db 1140 CCCAGCCCCACAGCCATGATCTCAACCTGTGGAGGGCGGCACTTCCCCAACGGCAAC 1199
Qy 874 LeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThr 893
Db 1200 CTCAGCCAGCTGGT-GCAGCAGTGGCTGGACTGGGAGCAGC--AGACGCTGGCTCTTCA 1255
Qy 894 ValSerGluAlaGluCys 899
Db 1256 GTGTGGAGGCTGAGTGC 1273
```

RESULT 8

```
US-09-833-381-1807
; Sequence 1807, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1807
; LENGTH: 1605
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1605)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1807
```

Alignment Scores:

Pred. No.:	4,72e-99	Length:	1605
Score:	1123.00	Matches:	254
Percent Similarity:	62.38%	Conservative:	81
Best local Similarity:	47.30%	Mismatches:	155
Query Match:	23.46%	Indels:	47
DB:	4	Gaps:	12

US-09-970-944-2 (1-899) x US-09-833-381-1807 (1-1605)

```
Qy 385 CysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSer---IleLeuThrSer 403
Db 19 TGCGGT-----GACTTCGACACAGACACTCACTGACTCATCTGCTGCTGACTGCTGT 69
Qy 404 GlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu----- 421
Db 70 GGTTCACCCCGCTCAACTTTAAGACGCGGACCAACCCGACGCTCCTACACGCC 129
Qy 422 ThrIleGlnProAspLeuSerThrThrThr---ThrTyrGlnGlySerLeuCysProArg 440
Db 130 TCTGTGCTCTCTGACTGACGCGCAGCGCGGCGCATCTACCGCGGACCCCGTGTATGCCCTG 189
Qy 441 GlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGly 460
Db 190 CAGGACTCC---ACGCACAAAATCCCATGACCAACTCTCTCTCTGTGAGCCCTTACCC 246
Qy 461 GlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAla----- 475
Db 247 AGCCTTAAGTCAAGGTCCACACCTCCAGCACACCGGGCTCTGGGCCAGGCTTGCAGAT 306
Qy 475 ----- 475
Db 307 GGGGCTGACCTGTGGGGGTCTTGCGGCTGGCACATACCTAGCGATTTCACCGGGGAC 366
Qy 476 GluGluPheValSerArgLeuSerThrGlnAsnTyrPhe-Arg-----SerLeuPr 492
Db 367 ACCCACTTCTGCACTCGCGCAGCGCCAGCTCGGTTCGCCAGCAGCTCTTGGGCGCTGCC 426
Qy 492 oArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIl 512
Db 427 CCGAGACCCAGGAGCAGCGTCAAGCGGACCTTTGGCTGCTGGGTGGGAGGCTCANCAT 486
Qy 512 eProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLysIleTyr 532
Db 487 CCCGGGACAGGGGTGAGCTTGTGTCGCCCAATGGAGCCATTCGCCAGGCAAGTTCTTA 546
Qy 532 rGlu-IleTyr-LeuThrLeuHisLysProGluAspValArgLeuProLeuAla---Gly 550
Db 547 CGAGATGTATTCTACTCATCAACAGGAGCAAGAGTACC---CTCCCGCTTTCAGAGGG 603
Qy 551 CysGlnThrLeuLeuSerProIleValSerCysGlyProGlyValLeuLeuThrArg 570
Db 604 ACCCANACAGTATTGAGCCCTCGTGACCTGTGAGCCACAGGCGCTCTGTGTGCGCG 663
Qy 571 ProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeuArg 590
```

```
Db 664 CCGCTCATCTCACCATGCCCACTGTGGCGAAGTCAGTGGCCGTCAGTCTTTCAG 723
Qy LeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGlu 610
Db 724 CTCAGACCCAGGCCACCAGGCCACTGG--GANGAGGTGGTGGACCTGGATGAGGAG 780
Qy AlaProSerHisLeuTyrTrpCysGlnLeuGluAlaSerAlaCysTyrValPheThrGlu 630
Db 781 ACCCTGAACACACCTCTGCTACTGCCAGCTGGAGCCAGGCCCTGTCCATCTCTGCTGGAC 840
Qy GlnLeuGlyArgPhe-AlaLeuValGlyGluAlaLeuSerValAlaAlaAlaAlaAla 650
Db 841 CAGCTGGGACCTCAGCGTGTTCAGGCGAGTCTTATCCGCTCAGCAGTCAAGCGGT 900
Qy ubLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTy 670
Db 901 CCAGCTGGCGGTCTTCGCCCGCCCTCTGCACCTCCCTGGAGTACAGCTCCGGGTCTA 960
Qy rCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuLysGlnLeuG 690
Db 961 CTGCTGGAGGACACCCCTGTAGCAGTGAAGGAGGTCTGGAGCTGGAGCGGACTCTGG 1020
Qy yGlyGlnLeuLysGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeuAr 710
Db 1021 CGGATACTTGGTGAGGACCGAAACCGCTAATGTTCAAGGACAGTACCACACCTGG 1080
Qy gLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyrG 730
Db 1081 CCTCTCCCTCCATGAGCTCCCTCCAGAGCTCCTCCAGAGCTCCTGCAAGATCTGGTGG 1260
Qy aGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPheA 790
Db 1261 TGAAGGGAGGAGGCCAGATATCCAGCTGCATACCACTCTGGCA--GAGACACCTGCTG 1317
Qy lacGluLeuLeuAlaLeuGluSerGluAlaGly-----ValProAlaLeuValGlyProS 808
Db 1318 GCTCCCTGGACACTCTGCTCTGCCCCCTGGCGAGCACTGTACCCACCGAGCTGGACCTT 1377
Qy exAlaPheLysIleProPheLeuLysGlnLysIleIleSerSerLeuAspProProC 828
Db 1378 ATGCTTCAAGATCCCACTGTCATCCCGCCAGAGATATGCAACAGCTAGATGCCCCCA 1437
Qy ysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuS 848
Db 1438 ACTCAGCGGGCAATGATGCTGGCGATGTAGCAGAGAAGCTCTCTATGGACCGGTACTGA 1497
Qy exPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgH 868
Db 1498 ATTACTTTGCCCAAGAGGAGGCCCGCCCGGNGTGATCTGGACCTCTGGAGAGCTCTGC 1557
Qy isPheProAsnGlyAsnLeuSerGlnLeuAlaAlaVal 881
Db 1558 AGCAGGACGATGGGACCTCAACAGGCTGCGAGTGCCTTG 1598
```

RESULT 9

```
US-08-253-155A-17/c
; Sequence 17, Application US/08253155A
; Patent No. 5691147
; GENERAL INFORMATION:
; APPLICANT: Gyuris, Jenio
; APPLICANT: Diaretta, Giulio
; TITLE OF INVENTION: CDK4 Binding Proteins
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/253,155A
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-028
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-253-155A-17
;
; Alignment Scores:
; Pred. No.: 8,7e-49 Length: 771
; Score: 600.50 Matches: 131
; Percent Similarity: 67.47% Conservative: 37
; Best Local Similarity: 52.61% Mismatches: 77
; Query Match: 12.54% Indels: 6
; DB: 1 Gaps: 2
```

```
US-09-970-944-2 (1-899) x US-08-253-155A-17 (1-771)
Qy 541 ProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSer 560
Db 769 CTTATGATGTGCCA---CATTATGCCTCTCCGAGACAGTATTGAGCCCTCGGTGACC 713
Qy 561 CysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGly 580
Db 712 TGTGGACCCACAGGCTCTGCTGTGCGCGCCGTCATCTCACCATGCCCCACTGTGCC 653
Qy 581 GluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrp 600
Db 652 GAAGTCAGTGGCCCGTGGACTGATCTTTCAGCTCAAGACCCAGGCCACCAGGCCACTGG 593
Qy 601 GluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTrpCysGlnLeu 620
Db 592 GAG---GAGTGGTGACCTGGATGAGGAGACCTCAACACACCTCTACTGCCAGCTG 536
Qy 621 GluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGlu 640
Db 535 GAGCCCGAGGCGCTGTACATCTCTGTGGACCACTGGGCGACCTCGTGTTCACGGGCGAG 476
Qy 641 AlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAlaCys 660
Db 475 TCTATTCCCGTTCAGAGCTCAAGCGGCTCCAGTGGCGCTTTCGCCCGCCCTCTGC 416
Qy 661 ThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLys 680
Db 415 ACCTCCCTGGAGTACAGACTCCGGGTCTACTGTCTGGAGGACAGGCGCTGTAGCACTGAAG 356
Qy 681 GluValValGlnLeuGluLysGlnLeuGlyGlnLeuLeuIleGlnLeuProArgValLeu 700
Db 355 GAGGTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCGCAACCGCTA 296
```

QY 701 HisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeu 720
 Db 295 ATGTTCAAGGACAGTAAACAACTCGCTCTCTC-CATGACCTCCCCATGCCCAT 237
 QY 721 TrpLysSerLysLeuValSerTyrGlnGluLeuProPheTyrHisIleTyrAsnGly 740
 Db 236 TGGAGGACAGCTGCTGGCAATACAGGAGATCCCTCTATCAATTGGAGTGA 177
 QY 741 ThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAsp 760
 Db 176 GCCAGAAGC--CTCCACTGCATTTCCCTCGAGAGGACAGGTTGGCTCCACAGAGC 119
 QY 761 LeuAlaCysLys-LeuTyrValTyrGlnVal-GluGlyAspGlyGlnSerPheSerIleA 780
 Db 118 TCAACCTGCAAGGATCTCGTGGGAAGTTGGAAAGGGGAGGACAGATATTCACAGTGC 59
 QY 780 snPheAsnIleThrLysAspThr 787
 Db 58 ATACCACTCTGGGAGGACACC 36
 RESULT 10
 US-09-262-537-48
 ; Sequence 48, Application US/09262537
 ; Patent No. 6479256
 ; GENERAL INFORMATION:
 ; APPLICANT: Haylick, Joel
 ; TITLE OF INVENTION: Lectomedin Materials and Methods
 ; FILE REFERENCE: 27866/35307
 ; CURRENT APPLICATION NUMBER: US/09/262,537
 ; CURRENT FILING DATE: 1999-03-04
 ; EARLIER APPLICATION NUMBER: 60/076,782
 ; EARLIER FILING DATE: 1998-03-04
 ; NUMBER OF SEQ ID NOS: 64
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 48
 ; LENGTH: 5749
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-262-537-48
 Alignment Scores:
 Pred. No.: Length: 5749
 Score: 267.50 Matches: 108
 Percent Similarity: 38.69% Conservative: 58
 Best Local Similarity: 25.17% Mismatches: 132
 Query Match: 5.59% Indels: 134
 DB: 4 Gaps: 21
 US-09-970-944-2 (1-899) x US-09-262-537-48 (1-5749)
 QY 30 ThrValAlaAsnProValProGly----- 37
 Db 901 AGCGTGGAAATCCAGATCCAAACCAATATAGCATTTACCTGAAATTTTCCAAAAAGGAC 960
 QY 38 -----AlaAsnProAspLeu-----ProHisPheLeuValGluPro 50
 Db 961 CTTAGTCTCTAACTTTTCACTCTGGCTTATCAGTTTGATCATTTTCCCAAGTATGCT 1020
 QY 51 GluAspValTyrIleValLysAsnLysProValLeuValCysLysAlaValProAla 70
 Db 1021 ATAAAGATCTTTTAAAGAAAGATCATTTCTATATGCACTCTGCAATTTCCCAAGTATGCT 1080
 QY 71 ThrGlnIlePhePheLysCysAsnGlyGluTyrValArgGlnValAspHisValIleGlu 90
 Db 1081 TTC---GTTTCTACAGTATGATAAAATTTTATT---CAATACCTCGAGTATTT--- 1131
 QY 91 ArgSerThrAspGlySerSerGlyGluProThrMetGluValArgIleAsnValSerArg 110
 Db 1132 -----CCAACCTAAATTT-CCCAGATTACAGAAAAAAGG 1163
 QY 111 GlnGlnValGluLys-----ValPheGlyLeuGluGlu----- 121
 Db 1164 GGAAGAGATCAGAAATCTTTTGGATTGTAATGAAACAAGGTACGCCCAAGCCA 1223

QY 122 TyrTrpCysGlnCysValAlaTyrSerSerSerGlyThrThrLysSerGlnLysAlaTyr 141
 Db 1224 GTTGTGGTGGCATGATATATGTTACTTGGTTGGAGAGCTGCTTAAATCATGAAATGGGAG 1283
 QY 142 IleArgIle-----AlaArgLeuArgLysAsnPhe----- 151
 Db 1284 AACAGATCATGTGGATCATGTATACAAATGCACCTGCCCTCAGCATTTGGGAGAGTG 1343
 QY 152 ---GluGlnGluProLeuAlaLysGluValSer-----Leu 162
 Db 1344 GGGGATCGACAGCAGCTCGCTGATTTTGTAAATAACGTGGTGTACCCCTGAATGAGCA 1403
 QY 163 GluGlnGlyIleValLeuPro-----CysArgProProGlu---GlyIleProProAla 179
 Db 1404 GACAGGGGCTGCCTGA-CCAGAGAGCTGCAACACCCCAAGCTGCAATCTTACACGGG 1462
 QY 180 GluValGluTyrLeuArgAsnGluAspLeuValAspProSerLeuAspPro-----Asn 197
 Db 1463 AGG-----CCAAGCGACCCCAAGAGAAT 1489
 QY 198 ValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArg----- 212
 Db 1490 TTGGA-ATGATGGGAGATCATACAAATTAAGTACGACGACCTCGATCTGTTTCATGAAAA 1548
 QY 213 -----LeuAlaAspThrAlaAsnTyr----- 219
 Db 1549 AGGTCCCTCAGGAACAAGCTGATCTGCTAAATTTATGGCAACAACCTGGTGAATCTGT 1608
 QY 219 ----- 219
 Db 1609 GTGGAAGAGTGTCCAGTGGAGCACATGTTGGTTACTTGTGTCAAGGGTCGCAGGTG 1668
 QY 220 -----ThrCysValAla-----LysAsnIleValAlaArgArg 230
 Db 1669 CGAACCCAGAACTTGTGTATACCTTTACGGGCACACTGCAGCGGCCCATTAAGAGAATCA 1728
 QY 231 ArgSerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTyrSerThrThrThrGlu 250
 Db 1729 AGGGTTGCATAACACTGCGCTCTGTCCAGTACAGGAGTATGGAGAAATGGTCACCA 1788
 QY 251 TrpSerValCysSerAlaSerCysGlyArgGlyTyrGlnLysArgSerArgSerCysThr 270
 Db 1789 TGGAGTTTATGTTTACATGTGTGAGGCCAAAGAACAAAGAACAGGTCATGCACA 1848
 QY 271 AsnProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThr 290
 Db 1849 -----CCTCCTCAGTATGGAGAGCGCGTGTGAAGGACCTGAAACACATCATAGCCT 1902
 QY 291 ValSerSerLeuLeuValSerValAspGlySerTyrTyrProTyrSerLysTyrSerAla 310
 Db 1903 TGTATATTTGCTCTTTGCCAGTTGATGACAGTGGCAAGATGGAGTTCGTGGAGCCAG 1962
 QY 311 CysGlyLeuAspCys-----ThrHisTyrArgSerArgGluCysSerAspProAla 327
 Db 1963 TGCTCAGTAACCTGCTCGAATGGGACTCAGCAGAGAGAGCGCGCAGTGCACT-----GCA 2016
 QY 328 ProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSer 347
 Db 2017 GCTGCCCATGGAGGCTCCGAATGCAGAGGGCCATCGGCAGAAAGCAGAGAGTGTATAAC 2076
 QY 348 AspLeuCysValHisSerAlaSerGly 356
 Db 2077 CCTGAATGT-----ACAGCCAATGT 2097
 RESULT 11
 US-08-808-982-4
 ; Sequence 4, Application US/08808982
 ; Patent No. 5939271
 ; GENERAL INFORMATION:
 ; APPLICANT: Tessier-Lavigne, Marc
 ; APPLICANT: Leonardo, E. David
 ; APPLICANT: Hink, Lindsay

APPLICANT: Masu, Masayuki
APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,982
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC96-217
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 305 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-808-982-4

Alignment Scores:
Pred. No.: Length: 305
Score: 253.00 Matches: 56
Percent Similarity: 71.43% Conservative: 14
Best Local Similarity: 57.14% Mismatches: 28
Query Match: 5.29% Indels: 2
DB: Gaps: 0

US-09-970-944-2 (1-899) x US-08-808-982-4 (1-305)

QY 612 ProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGln 631
Db 14 CCGAACACACCCCTGCTACTGAGCTGGAGCCCGAGCCCTGTAC-ATCCTGTGGACCAG 72
QY 632 LeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaAlaArgLeuLys 651
Db 73 CTGGGCACCTAGCTTTTCAGGGCGAGTCTTATCCCGCTCAGCAGTCAAGCGGCTCCAG 132
QY 652 LeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCys 671
Db 133 CTGGCCCGT-TTGGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCCTCCGGGTCTACTGC 191
QY 672 LeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGly 691
Db 192 CTGGAGGACAGCCTGTGTAGCAGTGAAGAGGTGCTGGAGCTGGAGCTGGAGCTCTGGGCGGA 251
QY 692 GlnLeuIleGlnProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709
Db 252 TACTTGGTGGAGAGCCGAAACCGCTAATGTTCAAGGACAGATTACCAACACCTT 305

RESULT 12
US-09-306-902A-4
Sequence 4, Application US/09306902A
Patent No. 6277585
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
Leonardo, E. David

Hink, Lindsay
Masu, Masayuki
Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/306,902A
FILING DATE: 07-May-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC96-217
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 305 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-306-902A-4

Alignment Scores:
Pred. No.: Length: 305
Score: 253.00 Matches: 56
Percent Similarity: 71.43% Conservative: 14
Best Local Similarity: 57.14% Mismatches: 28
Query Match: 5.29% Indels: 2
DB: Gaps: 0

US-09-970-944-2 (1-899) x US-09-306-902A-4 (1-305)

QY 612 ProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGln 631
Db 14 CCGAACACACCCCTGCTACTGAGCTGGAGCCCGAGCCCTGTAC-ATCCTGTGGACCAG 72
QY 632 LeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaAlaArgLeuLys 651
Db 73 CTGGGCACCTAGCTTTTCAGGGCGAGTCTTATCCCGCTCAGCAGTCAAGCGGCTCCAG 132
QY 652 LeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCys 671
Db 133 CTGGCCCGT-TTGGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCCTCCGGGTCTACTGC 191
QY 672 LeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGly 691
Db 192 CTGGAGGACAGCCTGTGTAGCAGTGAAGAGGTGCTGGAGCTGGAGCTGGAGCTCTGGGCGGA 251
QY 692 GlnLeuIleGlnProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709
Db 252 TACTTGGTGGAGAGCCGAAACCGCTAATGTTCAAGGACAGATTACCAACACCTT 305

RESULT 13
US-08-985-526-4
Sequence 4, Application US/08985526
Patent No. 6080728
GENERAL INFORMATION:

; APPLICANT: Mixson, James A
 ; TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA
 ; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
 ; TITLE OF INVENTION: THERAPY
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Connolly, Bove, Lodge, & Hutz
 ; STREET: 1220 Market Street, P.O. Box 2207
 ; CITY: Wilmington
 ; STATE: Delaware
 ; COUNTRY: U.S.A.
 ; ZIP: 19899
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/985,526
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/608,845
 ; FILING DATE: 16-JUL-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McMorrow Jr., Robert G
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (302) 658-9141
 ; TELEFAX: (302) 658-5613
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1326 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-985-526-4

Alignment Scores:
 Pred. No.: 2,826-14 Length: 1326
 Score: 250.00 Matches: 104
 Percent Similarity: 35.81% Conservative: 50
 Best Local Similarity: 24.19% Mismatches: 152
 Query Match: 5.22% Indels: 126
 DB: 3 Gaps: 17

US-09-970-944-2 (1-899) x US-08-985-526-4 (1-1326)
 QY 13 GlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSerAlaThrValAla 32
 DB 92 GGACTGTTGATAGTGCA-----CTGAGTGTCACTGTCAGA 127
 QY 33 AsnProValProGlyAlaAsn-----ProAspLeuLeuProHis 45
 DB 128 ACTCAGTTACCATCTGCAAAAGGTGCTCTGCCCATCATGCTGCTCCATGCCACAG 187
 QY 46 PheLeuValGlu----- 49
 DB 188 TTCCTGATGGAGATGCTGCTCTCGCTGTGTGGCCCGAGCTGCGGACGATGGCTGTGT 247
 QY 50 -----ProGluAspValTyrIleValLysAsnLysProValLeuValCysLys 66
 DB 248 CTCACATGTCCGAGTGGACCTCTGTTCTACGAGCTGTGGCAATGAATTCAGACGGCG 307
 QY 67 AlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArg---GlnVal 85
 DB 308 GCGGCTCTCTGGCAT-AGCCTCAACACCGATGTGAGGGCTCTCTCGTCCGACACGACG 366
 QY 86 AspHisValIleGluArgSerThrAspGlySerGlyGluProThrMetGluValArg 105
 DB 367 TGCCACATT----- 375
 QY 106 IleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrpCysGln 125

Db 376 -----CAGGAGTGTGCACAAAAGATTATAACAGGATGGTGGCTGGAGCCAC 420
 QY 126 CysValAlaLafTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAla 145
 Db 421 TGGTCCCGTGGTGCATCTTGTGTGACATGTGTGTGATGGTGTGATCACAAGGATC--- 477
 QY 146 ArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuLeuGlnIle 165
 Db 478 CGCTCTGCAACTCTCCAGCCCCCAGATGATGGAAGAACCTCTGGAAGCGGAGCGGG 537
 QY 165 yIleValLeuProCysArgProProGluGly----- 175
 Db 538 GAGACAAAGCCTGCAAGAAAGACGCTGCCCATCAATGAGGCTGGGGTCTTGGTCA 597
 QY 176 -----IleProProAlaGluValGluTrpLeuArgAsnGluAspLeuVa 190
 Db 598 CCATGGGACATCTGTTCTGTACCTGTGGAG---GAGGGGTACAGAAAAGCTAGTCGTCTC 654
 QY 190 LasPProSerLeuAsp----- 195
 Db 655 TGGCTCGACTCTAGAAATGACTGAAGAGAACAAAGAGTTGGCCCAATGAGCTGAGGCGGCT 714
 QY 196 -ProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAs 215
 Db 715 CCCCTATGCT---ATCACAAAGGAGTTCAGTACAGAAATAACGAGGAATGGACTGTT-GA 770
 QY 215 pThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgSerAlaSerAl 235
 Db 771 TAGCTGCACCT---GAGTGTCTACTGTCAAGACTAGTTACCATCTCAAAAAGGTGCTCTG 827
 QY 235 aAlaValIleValTyrValAsnGly----- 243
 Db 828 CCCCATCAATGCGCTCTCCATGCCACAGTTCCTGATGAGAAATGCTGCTCGCTGTGTG 887
 QY 244 -----GlyTrpSerThrTrpThrGluTrpSerValCysSerAl 256
 Db 888 GCCCAGCGACTCTGCGGAGCATGGTGGTCTCCATGGTCCGAGTGGACCTCCTGTTCTAC 947
 QY 256 sSerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAs 276
 Db 948 GAGCTGTGGCAATGAATTCAGCAGCGCGCGCTCTCTCGCATAGC-----CTCAA 998
 QY 276 nGlyGlyAlaPheCysGluGlnAsnValHisAspArgThrVal-----Se 292
 Db 999 CAACCGA-----TGTGAGGGCTCTCGTCCAGACACGACCTGCCACATTCAGAGTG 1052
 QY 292 rSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGl 312
 Db 1053 TGACAAAAGATTATAACAGGATGGTGGTGGAGCCACTGGTCCCGTGGTCTATCTGTTC 1112
 QY 312 yLeuAspCys-----ThrHisTrpArgSerArgGluCysSerAspProAlaProAr 329
 Db 1113 TGTGACATGTGTGTGATGTGTGATCACAAGGATCGGGCTCTGCAACTCTCCAGGCCCA 1172
 QY 329 gAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLe 349
 Db 1173 GATGAATGGAAACCTCTGTGAAGGGCGAAGCGGGGAGACCAAGAGCTGCAAGAAAGACGC 1232
 QY 349 uCys---ValHisSerAlaSerGlyPro 357
 Db 1233 CTGCCCCCATCAATGAGGCTGGGGTCTCT 1260

RESULT 14
 US-09-919-172-64
 ; Sequence 64, Application US/09919172
 ; Patent No. 6673545
 ; GENERAL INFORMATION:
 ; APPLICANT: Faris, Mary
 ; APPLICANT: Turner, Christopher M.
 ; TITLE OF INVENTION: PROSTATE CANCER MARKERS
 ; FILE REFERENCE: PA-0036 US
 ; CURRENT APPLICATION NUMBER: US/09/919,172
 ; CURRENT FILING DATE: 2001-07-30

```
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 64
; LENGTH: 7231
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673545 1399366.20
; NAME/KEY: unsure
; LOCATION: 5601, 5609, 7107
; OTHER INFORMATION: a, t, c, g, or other
US-09-919-172-64

Alignment Scores:
Pred. No.: 5,64e-13 Length: 7231
Score: 250.00 Matches: 48
Percent Similarity: 45.81% Conservative: 23
Best Local Similarity: 30.97% Mismatches: 66
Query Match: 5.22% Indels: 18
DB: 4 Gaps: 3

US-09-970-944-2 (1-899) x US-09-919-172-64 (1-7231)

QY 207 ValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsnIle 226
Db 1364 ATTCAGCAGCGGCGCTCTCGGATACGCTCAACACCGATGTGAGGGCTCTCGGTC 1423
QY 227 ValAlaArg-----ArgArgSerAlaSerAlaAlaValIle 238
Db 1424 CAGACAGGACCTGCCACATTCAGGAGTGTGACAGAGATTTAAACAG----- 1471
QY 239 ValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCys 258
Db 1472 -----GATGGTGGTGGAGCCACTGTCCTCCCGTGCATCTGTCTGTGACATGT 1522
QY 259 GlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGly 278
Db 1523 GGTGATGGTGTATCACAAGGATCCGGCTCTCAACTCTCCAGCCGCCAGATGAACGGG 1582
QY 279 AlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerVal 298
Db 1583 AAACCTGTGAAGCGAAGCGGGAGACCAAGCCCTCAAGAACGACCCCTGCCCATC 1642
QY 299 AspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCys----- 315
Db 1643 AATGAGGCTGGGGTCTTGTGTGTCACCATGGGACATCTGTCTGTACCTGTGGAGGAGG 1702
QY 316 ThrHisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluCys 335
Db 1703 GTACAGAACGTAGTCTCTGCAACACACCCACACCCAGTTTGGAGGCAAGGACTGC 1762
QY 336 GlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCys 350
Db 1763 GTTGTGTGTGTACAGAAACACAGATCTGCAACACAGCAGGACTGT 1807

RESULT 15
US-08-985-526-2
; Sequence 2, Application US/0898526
; Patent No. 6080728
; GENERAL INFORMATION:
; APPLICANT: Mixson, James A
; TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA
; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Connolly, Bove, Lodge, & Hutz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware
```

```
; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,526
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorow Jr., Robert G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 657 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-985-526-2

Alignment Scores:
Pred. No.: 2,24e-14 Length: 657
Score: 245.50 Matches: 49
Percent Similarity: 53.28% Conservative: 16
Best Local Similarity: 40.18% Mismatches: 44
Query Match: 5.13% Indels: 13
DB: 3 Gaps: 5

US-09-970-944-2 (1-899) x US-08-985-526-2 (1-657)
```

```
QY 244 GlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGln 263
Db 241 GCGTGGTCTCAAGTGTGGAGTGGACCTCTCTTCTACAGCTGTGGCAATGGAATTCAG 300
QY 264 LysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGly 283
Db 301 CAGCGCGCGCTCTCTGCGATAGC-----CTCAACACCGA-----TGTGAGGCG 345
QY 284 GlnAsnValHisAspArgThrVal-----SerSerLeuLeuValSerValAsp 299
Db 346 TCCTCGTCCAGACACGCGACCTGCCACATTCAGGAGTGTGACAAAGATTAAACAGGAT 405
QY 300 GlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCys-----Thr 316
Db 406 GGTGCTGGAGCCACTGGTCCCGGTGTCATCTTCTGTGACATGTGGTGTGATGTTG 465
QY 317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGln 336
Db 466 ATCAGAGATCCGGCTCTGCAACTCTCCAGCCCCCAGATGAATGGGAAACCCCTGTGAA 525
QY 337 GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCys---ValHisSerAlaSer 355
Db 526 GCGGAAGCGCGGAGACCAAGCCCTGCAAGAAAGACGCTGCCCATCAATGGAGCTGG 585
QY 356 GlyPro 357
Db 586 GGTCT 591
```

Search completed: October 6, 2004, 22:39:13
Job time : 210 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 6, 2004, 18:22:36 ; Search time 1039 Seconds
(without alignments)
4386.468 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRGMLPALLGIVLAAML.....AVAGLQPDAGLFTVSEAC 899

Scoring table:

	BLOSUM62	Xgapext	0.5
Ygapop 10.0			0.5
Fgapop 10.0			0.5
Fgapop 6.0			7.0
Delop 6.0			7.0

Searched: 3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09970944/runat_05102004_112011_1685/app_query.fasta_1.1095
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNIT5=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09970944 @CEN 1 1.622 @runat_05102004_112011_1685
-NCPU=6 -ICPU=3 -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRAD=10 -XGAPOP=0.5
-XGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

ALIGNMENTS

RESULT 1

US-09-970-944-1

; Sequence 1, Application US/09970944

; Publication No. US20030204052A1

; GENERAL INFORMATION:

; APPLICANT: Herrman, John L

; APPLICANT: Rastelli, Luca

; TITLE OF INVENTION: Shinkets, Richard A

; TITLE OF INVENTION: No. US20030204052A1e1 Proteins and Nucleic Acids Encoding Same and Antibodies Directed Against these Proteins

; FILE REFERENCE: 21402-138

; CURRENT APPLICATION NUMBER: US/09/970,944

; PRIOR FILING DATE: 2002-05-02

; PRIOR FILING DATE: 2000-10-04

; NUMBER OF SEQ IDS: 62

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 2881

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-970-944-1

1	4787	100.0	2881	11	US-09-970-944-1	Sequence 1, Appl
2	4698.5	98.2	2752	13	US-09-918-779-1	Sequence 1, Appl
3	4698.5	98.2	2752	17	US-10-624-932-1	Sequence 1, Appl
4	4545.5	95.0	2697	15	US-10-240-154-15	Sequence 15, Appl
5	4545.5	95.0	3014	10	US-09-933-261-1	Sequence 1, Appl
6	4545.5	95.0	3014	15	US-10-256-702-1	Sequence 1, Appl
7	4379.5	91.5	3580	17	US-10-311-623-13	Sequence 13, Appl
8	2694	56.3	1787	10	US-09-933-261-2	Sequence 2, Appl
9	2694	56.3	1787	15	US-10-256-702-2	Sequence 2, Appl
10	2490.5	52.0	2295	13	US-09-972-211-55	Sequence 55, Appl
11	2490.5	52.0	2295	13	US-10-096-625-55	Sequence 55, Appl
12	2484	51.9	2860	13	US-10-087-684-1	Sequence 1, Appl
13	2484	51.9	2860	13	US-10-087-684-3	Sequence 3, Appl
14	2484	51.9	2860	13	US-10-218-779-1	Sequence 1, Appl
15	2484	51.9	2860	13	US-10-218-779-3	Sequence 3, Appl
16	2479	51.8	3485	9	US-09-816-828-18	Sequence 18, Appl
17	2479	51.8	3485	13	US-10-147-493-145	Sequence 145, App
18	2479	51.8	3485	13	US-10-145-127-145	Sequence 145, App
19	2479	51.8	3485	13	US-10-160-503-145	Sequence 145, App
20	2479	51.8	3485	13	US-10-143-118-145	Sequence 145, App
21	2479	51.8	3485	13	US-10-144-993-145	Sequence 145, App
22	2479	51.8	3485	13	US-10-158-787-145	Sequence 145, App
23	2479	51.8	3485	13	US-10-140-024-145	Sequence 145, App
24	2479	51.8	3485	13	US-10-140-808-145	Sequence 145, App
25	2479	51.8	3485	13	US-10-152-405-145	Sequence 145, App
26	2479	51.8	3485	13	US-10-127-852A-145	Sequence 145, App
27	2479	51.8	3485	13	US-10-127-900A-145	Sequence 145, App
28	2479	51.8	3485	13	US-10-128-685A-145	Sequence 145, App
29	2479	51.8	3485	13	US-10-131-820A-145	Sequence 145, App
30	2479	51.8	3485	13	US-10-142-886-145	Sequence 145, App
31	2479	51.8	3485	13	US-10-146-728-145	Sequence 145, App
32	2479	51.8	3485	13	US-10-146-786-145	Sequence 145, App
33	2479	51.8	3485	13	US-10-147-499-145	Sequence 145, App
34	2479	51.8	3485	13	US-10-157-798-145	Sequence 145, App
35	2479	51.8	3485	15	US-10-028-072-145	Sequence 145, App
36	2479	51.8	3485	15	US-10-121-049-145	Sequence 145, App
37	2479	51.8	3485	15	US-10-123-904-145	Sequence 145, App
38	2479	51.8	3485	15	US-10-140-470-145	Sequence 145, App
39	2479	51.8	3485	15	US-10-175-746-145	Sequence 145, App
40	2479	51.8	3485	15	US-10-176-918-145	Sequence 145, App
41	2479	51.8	3485	15	US-10-176-921-145	Sequence 145, App
42	2479	51.8	3485	15	US-10-137-865-145	Sequence 145, App
43	2479	51.8	3485	15	US-10-140-474-145	Sequence 145, App
44	2479	51.8	3485	15	US-10-142-431-145	Sequence 145, App
45	2479	51.8	3485	15	US-10-143-114-145	Sequence 145, App

Alignment Scores:

Pred. No.: 0
 Score: 4787.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 11
 DB: 11

Length:	2881
Matches:	899
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0
US-09-970-944-2 (1-899) x US-09-970-944-1 (1-2881)	
Qy	1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db	87 ATGGCCGTCGGCCGCGCTGTGGCCACGCTCTCTGGGCATAGTCTCTCGCGCTTGGCTC 146
Qy	21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db	147 CGCGGCTCGGGTGGCCAGCAGATGCCACCTGGGCCAACCCAGTGGCTGGTGGCAACCG 206
Qy	41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db	207 GACCTGCTTCCCACTTCTCTGGTGAGCCCGAGGATGTGTACATCGTCAAGAACAAAGCCA 266
Qy	61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db	267 GTGCTGTGTGTGCAAGGCGCTGCCCGCCAGCAGATCTTCTTCAAGTGTCAACGGGGAG 326
Qy	81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db	327 TGGTGGCCAGGTGGACACCGTATCGAGCGCAGCACAGCGGGAGCGAGTGTGGAGCGG 386
Qy	101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db	387 ACCATGGAGGTCCGATTAATGTCTCAAGCGACGAGGTCTGAGAGAGTGTTCGGGGTGGAG 446
Qy	121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrLysSerGlnLysAla 140
Db	447 GAATACTCGTGCAGTGGTGATGGAGTCTCTCGGGCACCACCAAGAGTCAAGAGGCC 506
Qy	141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db	507 TACATCCGATAGCCAGATGGCGCAAGAACTTCGAGCAGGAGCGCTGGCCCAAGGAGTG 566
Qy	161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db	567 TCCCTGGAGCAGGGCATCGTGTGCTGCTCCGCTCCACCGAGGCGATCCCTCCAGCCGAG 626
Qy	181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db	627 GTGGAGTGGCTCCGGAACGAGGACCTGTGGTGGACCCGCTCTGGACCCCAATGTATACATC 686
Qy	201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
Db	687 ACAGCGGAGCACAGCTGTGGTGGACAGCGCGCTTGTGTGACACGGCCCAACTACACC 746
Qy	221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240
Db	747 TGGTGGCCCAAGAACATCGTGGACGTCGCGCAGCGCTCTCGCTGTGTCATCGCTAC 806
Qy	241 ValAsnGlyGlyTrpSerThrTrpGluTrpSerValCysSerAlaSerCysGlyArg 260
Db	807 GTGAACGGTGGTGGTGGCTGGACCGAGTGTCTCGTCTGGAGCGCCAGCTGTGGGCGC 866
Qy	261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyIleAlaPhe 280
Db	867 GGCTGGCAGAAACGAGCGGAGCTGCACCAACCCGCGCTCTCAACGGGGCGCTTTC 926
Qy	281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
Db	927 TGTGAGGGGAGAAATGTGCCATGACCGCACCGTCTCTCTCTGTGTCTGTGGAGCGG 986
Qy	301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db	987 AGCTGGAGCCCGTGGAGCAAGTGTGTGGCTGTGGGCTGGACTGGACCCACCTGGCGGAGC 1046

Qy	321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db	1047 CGTGAGTGCTCTGACCCAGCACCCCGCAAGAGGGAGGAGTGCCAGGCACTGACCTG 1106
Qy	341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db	1107 GACACCCGCAACTGTACAGTGACCTCTGTGTACACAGTGTCTTGTGGCCCTGAGGACGTG 1166
Qy	361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
Db	1167 GCCCTCTATGTGGCCCTCATCGCGTGGCGCTGTCTGCTGGTCTGCTGTCTGTCTC 1226
Qy	381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
Db	1227 ATCTCTGTTTATTCGCCGGAAGAGGAGGGCTGGACTCAGATGTGGCTGACTCTGCTCAT 1286
Qy	401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db	1287 CTCACCTCAGCTTCCAGCCCGTCAAGATCAAGCCAGCAAGAGCAGCAACCCCATCTG 1346
Qy	421 LeuThrIleGlnProAspLeuSerThrThrThrThrThrThrThrThrThrThrThr 440
Db	1347 CTCACCATCCAGCGGACCTCAGCACCAACCACTACCCAGGGCAGTCTCTGTCCCCGG 1406
Qy	441 GlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGly 460
Db	1407 CAGATGGGCCAGGCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT 1466
Qy	461 GlyGlyArgHisThrLeuHisSerSerProThrSerGluAlaGluGluPheValSer 480
Db	1467 GCGGCCGCCACACACACACACACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGCTCTCC 1526
Qy	481 ArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyr 500
Db	1527 CGCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGACCCAGCAACATGACCTAT 1586
Qy	501 GlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeu 520
Db	1587 GGGACCTTCAACTCTCTCGGGGCGGCTGATCATCCCTAATACAGGTATCAGCCCTCTCC 1646
Qy	521 IleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrIleuThrLeuHisLys 540
Db	1647 ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAGCTGCACAAG 1706
Qy	541 ProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSer 560
Db	1707 CCGGAAGACGTGAGGTTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1766
Qy	561 CysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGly 580
Db	1767 TGTGGACCCCTCTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1826
Qy	581 GluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrp 600
Db	1827 GAGCCAGCCCTGACAGCTGGAGCTCGCTCGCTCAAAAAGCAGTCTGTCGAGGCGAGCTG 1886
Qy	601 GluGlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeu 620
Db	1887 GAGCAGGATGTGTCGACCTCGGAGGAGGCGCTCCACCTCTTACTTACTGCCAGCTG 1946
Qy	621 GluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGlu 640
Db	1947 GAGGCCAGTGCCTGTCTACGCTTTCACCGAGCAGTGGCGCTTTCCTGCTGGTGGAGAG 2006
Qy	641 AlaLeuSerValAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCys 660
Db	2007 GGCCTCAGCGTGGCTGCCCGCAAGCGCTCAAGCTGCTTCTGTGTGGCGGCTGGCTGC 2066
Qy	661 ThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLys 680
Db	2067 ACCTCCTCGAGTACAACTCCGGGTCTACTGCTCTGATGACACCCACGATGACACTCAAG 2126

Qy	661	GluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuLeuGlnGluProArgValLeu	700
Db	2127	GAGGTGGTGCAGCTGGAGAACAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCCCTG	2186
Qy	701	HisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeu	720
Db	2187	CAC TTCAGGACAGTTACCCAACTGGCGCTATCCATCCACGATGTGCCAGCTCCCTG	2246
Qy	721	TrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleLeuTyrAsnGly	740
Db	2247	TGGAAGAGTAAGCTCCCTGTGCAGCTACCAGGAGATCCCTTTTATCACATCTGGAATGGC	2306
Qy	741	ThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAsp	760
Db	2307	ACGCAGCGGTACTTGCACTGCACCTTCACCGCTGGAGCGTGTACAGCCACGACCTAGTGAC	2366
Qy	761	LeuAlaCysLysLeuLeuTyrValTrpGlnValIcLgLyAspGlyGlnSerPheSerIleAsn	780
Db	2367	CTGGCCTGCAAGCTGTGGTGTGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAAC	2426
Qy	781	PheAsnIleThrLysAspThrArgPheAlaGlnLeuLeuAlaLeuGluSerGluAlaGly	800
Db	2427	TTCAACATCACCAAGGACACAAGGTITGTCTGAGCTGTGGCTCTTGGAGAGTGGAAGCGGG	2486
Qy	801	ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuLeuArgGlnLysIle	820
Db	2487	GTCCAGCCCTGTGGTGGCCCGAGTGCCCTTCAAGATCCCTTCTCATTCGGCAGAGAATA	2546
Qy	821	IleSerSerLeuAspPropProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLys	840
Db	2547	ATTTCAGCGCTGACCCACCGCTGTAGCGGGGTGCCGACTGCGGAGCTCTGGCCCCAGAA	2606
Qy	841	LeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIle	860
Db	2607	CTCCACTGGACAGCCATCTCAGCTCTTTGGCTTCCAAGCCCCAGCCCCACAGCCATGATC	2666
Qy	861	LeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla	880
Db	2667	CTCAACTGTGGGAGCGGGCACTTCCCAACGGCAACCTTCAGCCAGCTGGCTGCAGCA	2726
Qy	881	ValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys	899
Db	2727	GTGGCTGGACTGGGCCAGGCAGACGCTGGCCCTTTCACAGTGTTCGAGGCTGAGTGC	2783

RESULT 2

US-09-918-779-1

; Sequence 1, Application US/09918779

; Publication No. US20030064369A1

; GENERAL INFORMATION:

APPLICANT: Taupier, Raymond

APPLICANT: padigaru, Muralidhara

; APPLICANT: Rastelli, Luca

APPLICANT: Spaderna, Steven

APPLICANT: Shimkets, Richard

APPLICANT: Zerhusen, Bryan

APPLICANT: Spytek, Kimberly

APPLICANT: Shenoy, Suresh

APPLICANT: LI, LI

APPLICANT: Gusev, Vladimir
APPLICANT: Gusev, Vladimir

APPLICANT: Grosse, William
APPLICANT: Albrecht, John

APPLICANT: Alsobrook, John
APPLICANT: Ienley, Denise

APPLICANT: Lepley, Denise
APPLICANT: Burgess Catherine

APPLICANT: Burgess, Catherine
APPLICANT: Gerlach, Valerie

APPLICANT: GERLACH, Valerie
AFFILIANT: Ellerman Karen

APPLICANT: ELIERMAN, Karen
APPLICANT: MacDougall, John

APPLICANT: MacDougall, John
APPLICANT: Stone, David

APPLICANT: SCOTT, DAVID
; APPLICANT: SMITHSON, GLENDA

; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same

FILE OF INVENTION: NOVEL PROTEINS and NUCLEIC ACIDS Encoding same
; FILE REFERENCE: 21402-074 US

; CURRENT APPLICATION NUMBER: US/09/918.779

; CURRENT FILING DATE: 2001-07-30

1 ; PRIOR APPLICATION NUMBER: 60/221,409

2011

```

; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/222,840
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,752
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,762
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,770
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,769
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/225,146
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,392
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,470
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,697
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: 60/263,662
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 60/281,645
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2752
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-779-1

```

Alignment Scores:

Pred. No.:	0	Length:	2752
Score:	4698.50	Matches:	888
Percent Similarity:	98.89%	Conservative:	2
Best Local Similarity:	98.67%	Mismatches:	7
Query Match:	98.15%	Indels:	3
DB:	13	Gaps:	3

US-09-970-944-2 (1-899) x US-09-918-779-1 (1-2752)

QY	1	MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu	20
Db	46	ATGGGCGTCCGGCCGGCGCTGTGGCCAGCGCTCTGGGCATAGTCTCGCGCTTGGCTC	105
QY	21	ArgGlySerGlyAlaGlnIleSerAlaThrValAlaIleAsnProValProGlyAlaIleAsnPro	40
Db	106	CGGGCTCGGGTGCACAGAGTGCCACCCTGGGCCAACCCAGTGCCTGGTGTGCCAACCCG	165
QY	41	AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleVallysAsnLysPro	60
Db	166	GACCTTGCTTCCCCACTTCTCGTGGAGCCGAGGATGTGTACATCTCTCAAGAACCAAGCCA	225
QY	61	ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu	80
Db	226	GTGCTGCTTGTCGACAGCCCGTCCGCCACGAGATCTTCTCAAGTGCACACGGGGAG	285
QY	81	TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro	100
Db	286	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGCGGAGCAGTGGGCTGCC	345
QY	101	ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu	120
Db	346	ACCATTGGAGGTCCGCATTAAATGCTCAAGGCACAGGTCTCGAAGAGTGTTTCGGGGTGGAG	405
QY	121	GluTyrTrpCysGlnCysValAlaIleTrpSerSerSerGlyThrThrLysSerGlnLysAla	140
Db	406	GAATACTGGTCCAGTGGGTGGCATGGAGCTCTCTGGGCAACCCACAGAGTCTCAGAAGGCC	465
QY	141	TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnIleProLeuAlaLysGluVal	160
Db	466	TACATCCGCATTAGCCAGATTGCACAAGAACTTCGACAGAGCCGCTGGCCCAAGAGGTG	525

Qy 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db 526 TCCCTGGAGCAGGGCATCGTGTGCTCCCTGCTCCACCGAGGGCATCCCTCCAGCCGAG 585
Qy 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db 586 GTGGAGTGGCTCCGGAACAGAGACCTGGTGACCGCTCCCTGGAGCCCAATGTATACATC 645
Qy 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
Db 646 ACGCGGAGCAGACGCTGGTGGTGAGCAGGCCCGCTTGTGACACGGCCAACTACACC 705
Qy 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240
Db 706 TGGTGGCCAAAGAACATCGTGACAGCTGGCGCAGCGCTCCGCTGCTATCGTCTAC 765
Qy 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
Db 766 GTGAACGGTGGGTGTCGACGTGGACCGAGTGGTCCGCTGTCAGCGCCAGCTGTGGGCGC 825
Qy 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
Db 826 GGCTGGCAGAAACGGAGCGGAGCTGCACCAACCGCGGCTCTCAACGGGGGCGCTTTC 885
Qy 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
Db 886 TGTGAGGGGCAGAAATGTC---CAGAAACAGCGCTGGCCACCGCTGTGCCCATAGACGGC 942
Qy 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db 943 AGCTGAGCGCGTGGAGCAAGTGTGCGCTGTGGGCTGGAGTGCACCCACTGTGGGAGC 1002
Qy 321 A-gGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db 1003 CGTGAGTGTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTG 1062
Qy 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db 1063 GACACCGCAACTGTACCACTGACCTCTGTGTACACAGTGTCTGGCCCTGAGGACGTG 1122
Qy 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
Db 1123 GCCCTCTATGTGGGCTCATCGCGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTCCT 1182
Qy 381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
Db 1183 ATCCTCGTTATTGCCGGAAGAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1242
Qy 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db 1243 CTCACCTCAGGCTTCCAGGCCGTCAGCATCAAGCCCGAGCAAGCAGACACCCCATCTG 1302
Qy 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThrThr 439
Db 1303 CTCACCATCAGCCGAGCACTCAGCACACACACACACACACACACACACACACACACAC 1362
Qy 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db 1363 CGGCGAGATGGGCCCCAGCCCCAAGTTCCAGCTCACAATGGGCACTGTCTCAGCCCCCTG 1422
Qy 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
Db 1423 GGTGGGCGCGCCACACACTGCACACAGCTCTCCACACCTCTCAGGCGGAGGAGTTCGTC 1482
Qy 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db 1483 TCCGCGCTCTCCACCCAGAACTACTTCGCTCCCTGCTCCCTGCTCCCGAGCACCACATGACC 1542
Qy 500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db 1543 TATGGGACCTTCAACTTCTCCTGGGGGCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1602
Qy 520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539

Db 1603 CTCATCCCCCAGATGCCATACCCCGAGGGAGATCATGAGATCTACTCTCACGCTGCAC 1662
Qy 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db 1663 AAGCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTAGAGCCCTGCTGAGTCCCATCGTT 1722
Qy 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db 1723 AGCTGTGACCCCTGGCGTCTGCTACCCCGGCCAGTCATCTGGCTATGAGCACCTGT 1782
Qy 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db 1783 GGGGAGCCCGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGCTCGTGGAGGGCAGC 1842
Qy 600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
Db 1843 TGGGAG---GATGTGCTGACCTGGCGGAGGAGCGCCCTCCACCTCTACTACTGCCAG 1899
Qy 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db 1900 CTGGAGGCCAGTGCCTGCTACGTCTTACCCGAGCAGCTGGCGCGCTTGGCCCTGGTGGGA 1959
Qy 640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
Db 1960 GAGGCCCTCAGCTGGCTGCCCGCAAGCGCTCAAGCTGCTTCTCTTGGCGCGGTGGCC 2019
Qy 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db 2020 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTCGCTGCATGACACCCAGCATGCATC 2079
Qy 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
Db 2080 AAGGAGGTGTCACGTGGAGAGCAGCTGGGGGGACAGCTGATCCAGAGGCCACCGGTC 2139
Qy 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db 2140 CTGCACTTCAAGSACAGTTACCAACACCTGCGCCTATCCATCCAGATGCGCCAGCTCC 2199
Qy 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db 2200 CTGTGGAGAGTAAGCTCCTTGTACGTACCAAGAGATCCCTTTTATCACATCTCGAAT 2259
Qy 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db 2260 GGCAGCAGCGGTACTTGCATTCACCTTACCTTGAGCGGTGTGAGCCCCAGCACTAGT 2319
Qy 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db 2320 GACCTGGCTGCAAGCTGTGGGTGGCAGGTGGAGGGCGAGCGGCAGAGCTTCAGCATC 2379
Qy 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db 2380 AACTTCAAGATCACAAGGACACAAGTTTGTGTAGCTGCTGGCTCTGGAGAGTGAAGCG 2439
Qy 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db 2440 GGGGTCCCAGCCCTGGTGGCCCCAGTGTCTTCAAGATCCCCTTCTCATTCGGCAGAG 2499
Qy 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db 2500 ATAATTTCCAGCTGACCCACCCCTAGCGGGGTGGCGACTCTGGCGGACTCTGGGCCAG 2559
Qy 840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db 2560 AAACTCCACTGACAGCCATCTCAGCTCTTTTGGCTTCCAAAGCCCGCCAGCCAGCCATG 2619
Qy 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db 2620 ATCCTCAACTGTGGAGGGCGCGCACTTCCCACCGCAACCTCAGCAGCTGGCTGCA 2679
Qy 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899

Db	2680	GCAGTGGCTGGACTGGGCCAGCAGAGCTGGCCCTCTTCACAGTGTCCGAGGCTGAGTGC	2739
RESULT 3			
US-10-624-932-1			
; Sequence 1, Application US/10624932			
; Publication NO. US20040096877A1			
; GENERAL INFORMATION:			
; APPLICANT: Taupier, Raymond			
; APPLICANT: Padigar, Muralidhara			
; APPLICANT: Rastelli, Luca			
; APPLICANT: Spaderma, Steven			
; APPLICANT: Shimkets, Richard			
; APPLICANT: Zerhusen, Bryan			
; APPLICANT: Spytek, Kimberly			
; APPLICANT: Shenoy, Suresh			
; APPLICANT: Li, Li			
; APPLICANT: Gusev, Vladimir			
; APPLICANT: Grosse, William			
; APPLICANT: Alsobrook, John			
; APPLICANT: Lepley, Denise			
; APPLICANT: Burgess, Catherine			
; APPLICANT: Gerlach, Valerie			
; APPLICANT: Ellerman, Karen			
; APPLICANT: MacDougall, John			
; APPLICANT: Stone, David			
; APPLICANT: Smithson, Glennda			
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same			
; FILE REFERENCE: 21402-074 US			
; CURRENT APPLICATION NUMBER: US/10/624,932			
; CURRENT FILING DATE: 2003-07-21			
; PRIOR APPLICATION NUMBER: 09/918,779			
; PRIOR FILING DATE: 2001-07-03			
; PRIOR APPLICATION NUMBER: 60/221,409			
; PRIOR FILING DATE: 2000-07-28			
; PRIOR APPLICATION NUMBER: 60/222,840			
; PRIOR FILING DATE: 2000-08-04			
; PRIOR APPLICATION NUMBER: 60/223,752			
; PRIOR FILING DATE: 2000-08-08			
; PRIOR APPLICATION NUMBER: 60/223,762			
; PRIOR FILING DATE: 2000-08-08			
; PRIOR APPLICATION NUMBER: 60/223,770			
; PRIOR FILING DATE: 2000-08-08			
; PRIOR APPLICATION NUMBER: 60/223,769			
; PRIOR FILING DATE: 2000-08-08			
; PRIOR APPLICATION NUMBER: 60/225,146			
; PRIOR FILING DATE: 2000-08-14			
; PRIOR APPLICATION NUMBER: 60/225,392			
; PRIOR FILING DATE: 2000-08-15			
; PRIOR APPLICATION NUMBER: 60/225,470			
; PRIOR FILING DATE: 2000-08-15			
; Remaining Prior Application data removed - See File Wrapper or PALM.			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 1			
; LENGTH: 2752			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-624-932-1			
Alignment Scores:			
Pred. No.:	0	Length:	2752
Score:	4698.50	Matches:	888
Percent Similarity:	98.89%	Conservative:	2
Best Local Similarity:	98.67%	Mismatches:	7
Query Match:	98.15%	Indels:	3
DB:	17	Gaps:	3
US-09-970-944-2 (1-899) x US-10-624-932-1 (1-2752)			
Qy	1	MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu	20
Db	46	ATGGCCGTCCGCGCCGCTGTGGCAGCGCTCTCTGGGCATAGTCTCGCGCTTGGCTC	105

Qy	21	ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro	40
Db	106	CGCGGCTCGGTGCCCGCAGCAGTCCACCGTGGCCACCCAGTGCCTGTGTCACCCG	165
Qy	41	AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro	60
Db	166	GACCTGCTTCCCACTTCTCTGTGGAGCCCGAGGATGTGTACATCTGTCAAGAACAGCCA	225
Qy	61	ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu	80
Db	226	GTGCTGCTTGTGTCAAGGCCGTGCCGCCACCGAGATCTTCTTCAAGTGTCAACGGGGAG	285
Qy	81	TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro	100
Db	286	TGGGTGCGCCAGGTGGACCACTGATCGAGCGCAGCAGACGCGGAGCTGGGCTGCCCC	345
Qy	101	ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu	120
Db	346	ACCATGGAGGTCCGCATTAAATGTCTCAGGCCAGCAGGTGAGAAGGTGTTCGGGCTGGAG	405
Qy	121	GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla	140
Db	406	GAATACTGCTGCCAGTGGCTGTCATGGAGCTCTCGGGCACCACCAAGAGTCAGAAGGCC	465
Qy	141	TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal	160
Db	466	TACATCCGATAGCCAGATTGGCAAGAACTTCGAGCAGAGCCGCTGCCAAGGAGGTG	525
Qy	161	SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu	180
Db	526	TCCCTGGAGCAGGGCATCGTGTGCTGCCCTGCCCTCACCGAGGGCATCCCTCCAGCCGAG	585
Qy	181	ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle	200
Db	586	GTGGAGTGGCTCCGGACAGGAGACCTGGTGGACCCGCTCCCTGGACCCCAATGTATACATC	645
Qy	201	ThrArgGluHisSerLeuValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr	220
Db	646	ACGCGGAGACAGCCTGGTGGCAGAGCCCGCTGTGTGACACGCGCCCACTACACC	705
Qy	221	CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr	240
Db	706	TGCGTGGCCCAAGAACATCGTGGCAGTCCCGCAGCGGCTCCGCTGCTCATCGTCTAC	765
Qy	241	ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg	260
Db	766	GTGAACGGTGGTGTGTGACCTGGACCGAGTGGTCCGTCTGCGCGCCAGCTGTGGCGCC	825
Qy	261	GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe	280
Db	826	GGCTGGCAAAACGGAGCCGGAGCTGCACCAACCCGGCGCTCTCAACGGGGCGCTTTC	885
Qy	281	CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly	300
Db	886	TGTGAGGGGCAGAAATGTC--CAGAAAACAGCCTGGCGCCACCTGTGCCAGTAGACGGC	942
Qy	301	SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer	320
Db	943	AGCTGGAGCCCGTGAGCAAGTGTGCGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	1002
Qy	321	ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu	340
Db	1003	CGTAGTGCTCTGACCCAGCACCCCGCAACCGAGGGAGGAGTGCCAGGGCAGCTGACCTG	1062
Qy	341	AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal	360
Db	1063	GACACCCGCAACTGTACCACTGACCTCTGTGTACACAGTGTCTTGGCCCTGAGGACGTG	1122
Qy	361	AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuValLeu	380
Db	1123	GCCCTCATGTGGGCTCATCGCGCTGGCGCTCTGCCTGGTCTCTGCTGCTGCTGCTC	1182
Qy	381	IleLeuValTyrCysArgLysLysGlyLeuAspSerAspValAlaAlaAspSerIle	400

```
Db 1183 ATCTCTCGTTTATGCGGAGAGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1242
QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db 1243 CTCACCTCAGGCTTCAGGCCCTCAGCATCAAGCCGAGCAAGAGCAGACACCCCATCTG 1302
QY 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlnSerLeuCysPro 439
Db 1303 CTCACCATCAGCGGACCTCAGCACCAACACCACTACAGGGCAGTCTCTGTCCC 1362
QY 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db 1363 CGGCGAGGATGGGCCCGGCCCCAAGTTCCAGCTCACAATGGGCACCTGCTCAGCCCCCTG 1422
QY 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluPheVal 479
Db 1423 GGTGGGGCGGCGCACACACTGCACACAGCTCTCCACCTCTGAGGCGGAGGATCGTC 1482
QY 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db 1483 TCCCGCCTCTCCACCGAGAACTACTTCCGCTCCCTGCCCGGAGGACCAACATGACC 1542
QY 500 TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db 1543 TATGGGACCTTCAACTTCTCGGGGCGGCGTATGATCCCTAATACAGGTATCAGCCTC 1602
QY 520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
Db 1603 CTCATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCAGCTGCAC 1662
QY 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db 1663 AAGCGGAAGAGCTGAGGTGGCTTGCCTTAGCTGGTGTGCAGACCTGTGAGTCCATCGT 1722
QY 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db 1723 AGCTGTGGACCCCTCGGCTGCTGCTCACCGGCCAGTCTATCTGGCTATGACCACTGT 1782
QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysGlnSerCysGluGlySer 599
Db 1783 GGGGAGCCCGAGCCTCAGAGCTGAGCCTCGGCTCAAAAAGCAGTGTGCGAGGGCAGC 1842
QY 600 TrpGluGlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGln 619
Db 1843 TGGGAG---GATGTGCTGCACCTGGCGGAGGAGCGCCCTCCACACCTCTACTACTCCAG 1899
QY 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db 1900 CTGGAGGCCAGTCCCTGCTACGTTCTACCGAGCAGCTGGGCCGCTTGGCCCTGGTGGGA 1959
QY 640 GluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
Db 1960 GAGGCCCTCAGCGTGGCTGGCGCAAGCGCTCAAGCTGCTTCTGTTGGCGCGGTGGCC 2019
QY 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db 2020 TGCACTTCCCTCAGTACAAATCCGCGTCTACTGCTGCATGACACCCACGATGACATC 2079
QY 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
Db 2080 AAGAGGTGTGCTGGAGAGCAGCTGGGGGACAGCTGATCCAGGAGCACCGGTC 2139
QY 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db 2140 CTGCACTTCAAGGACAGTTACCAACCTCGCGCTATCCATCCACGATGTGCCAGCTCC 2199
QY 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db 2200 CTGTGAAGAGTAAAGCTCTTGTGCTACGAGAGATCCCTTTTATCATCTGGAA 2259
QY 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
```

```
Db 2260 GGCAGCGAGCGGTACTTGCATGTCACCTTACCCTGGAGGTGTGCAGCCCGACACTAGT 2319
QY 760 AspLeuAlaCysLysLeuTyrValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db 2320 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGAGAGCTTTCAGCATC 2379
QY 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db 2380 AACTTCAATCATCACCAAGGACCAAGGTTTGTGTAGCTGTCTGGCTCTGGAGAGTGAAGCG 2439
QY 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db 2440 GGGGTCCAGCCCTGGTGGGCCCGAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAAG 2499
QY 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db 2500 ATAAATTCAGCCTGAGCCACCCCTGTAGCGGGGTGCCGACTGGCGGACTCTGGGCCAG 2559
QY 840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db 2560 AAATCCACCTGGAGCAGCATCTCAGCTTCTTTGGCTCCAGCCCGACCCCGACCCATG 2619
QY 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAla 879
Db 2620 ATCTTCAACTGTGGAGGGCGGCACATTCGCCAACGGCAACCTCAGCCAGCTGGTGCA 2679
QY 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 2680 GCAGTGGCTGGACTGGCCAGCAGACGCTGGCCCTTTCACAGTGTGGAGGCTGAGTGC 2739

RESULT 4
US-10-240-154-15
; Sequence 15, Application US/10240154
; Publication NO. US20030175741A1
; GENERAL INFORMATION:
; APPLICANT: Cochran et al.
; TITLE OF INVENTION: SCHIZOPHRENIA RELATED GENES
; FILE REFERENCE: CKPW-P01-006
; CURRENT APPLICATION NUMBER: US/10/240,154
; CURRENT FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: PCT/GB01/01486
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
; LENGTH: 2697
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2697)
US-10-240-154-15

Alignment Scores:
Pred. No.: 0 Length: 2697
Score: 4545.50 Matches: 852
Percent Similarity: 96.78% Conservative: 19
Best Local Similarity: 94.67% Mismatches: 26
Query Match: 94.96% Indels: 3
DB: 15 Gaps: 3

US-09-970-944-2 (1-899) x US-10-240-154-15 (1-2697)

QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyValLeuAlaAlaTrpLeu 20
Db 1 ATGGCCGCTCGGCCCGGCGCTGGCCAGTGCCTGGGCATAGTCTCGCCGCTGCGCTT 60
QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db 61 CGTGGTTCGGGTGCCCGAGCAGAGTGCACGGTGGCCCAATCCAGTGCCTGGTGCACACCCC 120
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
```


Db 121 GACCTGCTGCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAGAAACAAGCCG 180
Qy 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db 181 GTGTTGTTGGTGGCAAGGCTGGCTGCACCCACAGATCTTCTCAAGTGCATGGGGAA 240
Qy 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db 241 TGGGTCGCCAGGTCGATCACGTAATTAACGACGACCCACAGCAGCAGCGGATGGCA 300
Qy 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db 301 ACCATGGAGTCCGATCAACGATATCGAGCAGCAGGTAGAGAAAGTGTGGGCTGGAG 360
Qy 121 GluTyrTrpCysGlnCysValAlaTyrSerSerSerGlyThrThrLysSerGlnLysAla 140
Db 361 GAATACGTGGTCAGTGTGGCATGGAGCTCTCGGGTACCACCAAAAGTCAGAAAGGCC 420
Qy 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db 421 TACATCCGATTCCTATTGGCAGAACTTTGAGCAGAGCCACTGGCCAGGAAGTG 480
Qy 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db 481 TCACCTGGAGCAAGGCTTGTACTACCTTGTGCGCCGCCAGAGGAATCCGCCAGCTGAG 540
Qy 181 ValGluTyrLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db 541 GTGGAGTGGCTCGAAATGAGGACCTCTGGGACCCCTCCCTCGATCCCAATGTGTACATC 600
Qy 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
Db 601 ACGCGGAGCAGACCTAGTCTGTGGTCAGGCCCGCTGGCGCAGCAGCCCACTACACC 660
Qy 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240
Db 661 TGTGTGGCCAGAACATCTAGCCCGTGCAGGACCCCTCTGCAGCGGTCAATGTTTAT 720
Qy 241 ValAsnGlyGlyTyrSerThrThrThrGluTyrSerValCysSerAlaSerCysGlyArg 260
Db 721 GTGAACGGTGGTGGTCAGCGTGGAGTGTAGTGGTCCGCTGCAGCGCCAGCTGTGGCGGT 780
Qy 261 GlyTyrGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
Db 781 GCGTGGCAGAACGGAGCGGAGCTGCACCAACCCGGACCTCTCAACGGGGCGCCCTTC 840
Qy 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
Db 841 TGTGAGGGGCGAATGTC---CAGAAAAACAGCCTGGCCACTCTGTGCCAGTGGATGGG 897
Qy 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db 898 AGCTGGAGTCTGTGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957
Qy 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGlnGlyThrAspLeu 340
Db 958 CGCGAGTCTGTGACCCAGCACCCCGCAATGAGGTGAGAGTGTGGGGTGTCTGACCTG 1017
Qy 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db 1018 GACACCCGCAACTGTATACAGTACCTCTGCCTGCACACCGCTCTTGTGCCCGAGGACGTG 1077
Qy 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
Db 1078 GCTCTCATATCGGCTTGTGCTGTGGCTGTGGCTCTTCTTCTGTGTGGTGGCCCTT 1137
Qy 381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
Db 1138 GGACTCATTTACTGTGCAAGAGAGAGGGCTGGACTCGATGTGGCGGACTCGTCCATC 1197
Qy 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db 1198 CTCACCTCGGGGTTCCAGCCTGTACGATCAAGCCCGCAGCAAGCAGACACCCACCTG 1257

Qy 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrGlnGlySerLeuCysPro 439
Db 1258 CTCACCATCCAGCAGACCTCAGCACCACTACCACTACCACTACCACTACCACTACCACT 1317
Qy 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db 1318 AGCAGGATGGACCCAGCCCAAGTTCAGCTCTCTAATAGGTCACTGTCTACGCCACTG 1377
Qy 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
Db 1378 GGGAGTGGCGCGCATACGTTGGCACCACTCACCACCTCTCAGGCTCAGGACTTCGTC 1437
Qy 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db 1438 TCCCGCTCTCCACCAAACTACTTTCCTTCCCTGCCCGCGGACACGACCAATGGCC 1497
Qy 500 TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db 1498 TAGGGACCTTCACTTCTCGGGGCCCGCTGATGATCCCTAATACGGGGATCAGCCTC 1557
Qy 520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
Db 1558 CTCATACCCCGGATGCCATCCCGGAGGAAGATCTACAGATCTACCTCAGCTGCAC 1617
Qy 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db 1618 AAGCCAGAGAGCTGAGGTTGCCCTAGCTGCTGTCTCAGACCTGTCTGAGTCCAGTCGTT 1677
Qy 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db 1678 AGCTGTGGGCCCCAGAGTCTCTCTCAGCGGCGAGTCTCTTGCATGATGACCACTGT 1737
Qy 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db 1738 GGAGAGCCAGCCCTGACAGCTGGAGTCTGCGCTCNAAGAGCAGTCTCTCGAGGGCAGT 1797
Qy 600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrCysGln 619
Db 1798 TGGGAG---GATGTCTGCACCTTGTGAGGAGTCACTCTCCACCTCTACTACTGCCAG 1854
Qy 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db 1855 CTGGAGGCCGGGGCTGTCTATGTCTTTCAGGAGAGCTGGGGCGCTTGGCCCTGTAGGA 1914
Qy 640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla 659
Db 1915 GAGGCCCTCAGCTGGCTGCCACCAAGCGCTCAGGCTCTCTCTTTGCTCCCGTGGCC 1974
Qy 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db 1975 TGTAGTCCCTTGAGTACACATCCAGTGTACTGCCTACAGCAGACACCCAGCAGCTCTC 2034
Qy 680 LysGluValValGlnLeuLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
Db 2035 AAGAGGTGTGAGCTGGAGAGCAGTGTGTGACAGCTGTATCCCTCAGGAGCTCGCGTC 2094
Qy 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db 2095 CTGCACCTTCAAAGACAGTTACCAACACCTAGCTCTCTCCATCCAGCAGCTGCCAGCTCC 2154
Qy 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db 2155 CTGTGGAAGAGCAAGTACTTGTCTCAGCTACCGAGGAGTCCCTTTTACCATCTCGAAC 2214
Qy 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db 2215 GGCACCCAGAGTATCTGCATCTGCACCTTCACTCTGGAGCGCATCAACCGCAGCACCAGC 2274
Qy 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db 2275 GACCTGGGCTTCAAGGTTGGGTGTGGCAGGTGGAGGGAGATGGGAGAGCTTCAACATC 2334

QY 780 AsnPheAsnIleThrIysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
 |||||
 Db 2335 AACITTCACATCACTAAGACACAGGTTTGCTGAATTGTGGCTCTGGAGAGTGAAGG 2394
 |||||
 QY 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
 |||||
 Db 2395 GGGTCCAGCCCTGGTGGCCCCAGTGCCTTCAGATCCCTTCTCAATTGGCAAAAG 2454
 |||||
 QY 820 IleLeuSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
 |||||
 Db 2455 ATCATCGCAGCTGGACCAACCCCTGCAGCGGGCGCGCACTGGAGAACTCTAGCCAG 2514
 |||||
 QY 840 LysLeuHisLeuAspSerHisLeuSerPheAlaSerLysProSerProThrAlaMet 859
 |||||
 Db 2515 AAATTCACCTGACAGCCATCTAGCTTCTTGGCCCTCAAGCCCGCCCTACAGCCATG 2574
 |||||
 QY 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
 |||||
 Db 2575 ATCTCAACTATGGAGGACGCGCACTTCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634
 |||||
 QY 880 AlaValAlaGlyLeuGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
 |||||
 Db 2635 GCTGTGGCCGACTGGGCCAACACAGATGCTGGCCTCTTCACGGTGTGGAGGCCAGTGT 2694
 |||||

RESULT 5

US-09-933-261-1
 ; Sequence 1, Application US/09933261
 ; Publication No. US20030040046A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tessier-Lavigne, Marc
 ; Leonardo, E. David
 ; Hink, Lindsay
 ; Masu, Masayuki
 ; Kazuko, Keino-Masu
 ; TITLE OF INVENTION: Netrin Receptors
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
 ; STREET: 268 BUSH STREET, SUITE 3200
 ; CITY: SAN FRANCISCO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/933,261
 ; FILING DATE: 20-Aug-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/808,982
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OSMAN, RICHARD A
 ; REGISTRATION NUMBER: 36,627
 ; REFERENCE/DOCKET NUMBER: UC96-217
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 343-4341
 ; TELEFAX: (415) 343-4342
 ; INFORMATION FOR SEQ ID NO. 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3014 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 ; US-09-933-261-1
 ; Alignment Scores:

Pred. No.: 0 Length: 3014
 Score: 4545.50 Matches: 852
 Percent Similarity: 96.78% Conservative: 19
 Best Local Similarity: 94.67% Mismatches: 26
 Query Match: 94.96% Indels: 3
 DB: 10 Gaps: 3
 US-09-970-944-2 (1-899) x US-09-933-261-1 (1-3014)
 QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaLafTrpLeu 20
 |||||
 Db 1 ATGGCGCTCCGGCCCGGCTGTGGCCAGTGTCTCTGGCATAGTCTCGCCGCTGGCTT 60
 |||||
 QY 21 ArgGlySerGlyAlaGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
 |||||
 Db 61 CCGTGTGGGTGGCCAGCAGAGTGCACGGTGGCCAAATCCAGTGGCCGCTGCAACCCC 120
 |||||
 QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTrileValLysAsnLysPro 60
 |||||
 Db 121 GACCTGCTGCCCCACTTCTCTGGTAGAGCTGTAGGAGCTGTACATTGTCAAGAAACAAGCCG 180
 |||||
 QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
 |||||
 Db 181 GTGTGTGTGGTGTCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGTCAATGGGGA 240
 |||||
 QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
 |||||
 Db 241 TGGGTCCGCCAGGTGCATCACGTAATTGAACGAGCAGCAGCAGCAGCAGCGGATGGCA 300
 |||||
 QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
 |||||
 Db 301 ACCATGGAGTCCGTATCAACGTATCGAGCGAGCAGTAGAGAAAGTGTGGTGGTGGAG 360
 |||||
 QY 121 GluTrpTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
 |||||
 Db 361 GAATACTGTGGCAGTGTGTGGCATGGAGTCTCTCGGTACCACCAAAAATCTCAGAGGCC 420
 |||||
 QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
 |||||
 Db 421 TACATCCGATTGCCATTATTGGCAAGAACTTTGAGCAGGAGCCATGTGCCAAGGAAGTG 480
 |||||
 QY 161 SerLeuGluGlnGlyIleValLeuProCysArgProGluGlyIleProAlaGlu 180
 |||||
 Db 481 TCACTGGAGCAAGGCATTGTACTACTTGTGCCCCCAGAGGAATCCCCCAGCTGAG 540
 |||||
 QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
 |||||
 Db 541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCCGATCCCAATGTGTACATC 600
 |||||
 QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
 |||||
 Db 601 ACGGGGAGCAGAGCTGTGTGGCTGTAGGCCCGCCCTGGCCGACAGCGGCAACTACACC 660
 |||||
 QY 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaValIleValTyr 240
 |||||
 Db 661 TGTGTGCCCAAGAACATCGTAGCCCTGCGCGAAGCACCTCTCGAGCGGTCAATTGTTAT 720
 |||||
 QY 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerGlyArg 260
 |||||
 Db 721 GTGAACGGTGGGTGGTGCAGCTGAGTGTGTGCTGTGCAGCGCCACTCTGTGGCGGT 780
 |||||
 QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
 |||||
 Db 781 GGCTGGCAGNAACGGAGCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGCGCCTTC 840
 |||||
 QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
 |||||
 Db 841 TGTGAGGGGAGAAATGTC---CAGAAAAACAGCCCTGGCCCACTCTGTGCCAGTGGATGGG 897
 |||||
 QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
 |||||
 Db 898 AGCTGGAGTTCGTGGAGTAAGTGGTGTGGCTGTGGCTGTGCTGCACCCCACTGTGGCGAGC 957
 |||||

321 ArgGluCysSerAspProAlaProArgAsnGlyGluGluCysGlnGlyThrAspLeu 340
Db CGCGAGTGTCTGACCCAGCACCCCGAATGAGGTGAGAGTGTGGGGTGTGACCTG 1017
341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db GACACCGCAACTGTACAGTGTGCTGCTGCACACCGCTCTTGGCCCGAGGACGTG 1077
361 AlaLeuValGlyLeuLeuAlaValAlaValCysLeuValLeuLeuValLeu 380
Db GCTCTCTACATCGCCCTGTGGCTGTGGCTGTGCTCTTCTGTGTGGCCCTT 1137
381 IleLeuValTyrCysArgLysGlyGluLeuAspSerAspValAlaAspSerSerIle 400
Db GGACTCATTTACTGTGCGAAGAAGAGGGCTGGAATCCGATGTGGCCGACTGCTCCATC 1197
401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db CTCACCTCGGGCTTCCAGCCTGTGACGATCAAGCCAGCCAGCAAGACACACCCCACTG 1257
421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
Db CTCACATCCAGCCAGACCTCAGCACCAACCACTTACCCTACAGGGCAGTCTATGTTCG 1317
440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db AGGCAGGATGGACCCAGCCCAAGTTCAGCTCTCTAATGGTCACTGTCTAGCCCACTG 1377
460 GlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluPheVal 479
Db GGGAGTGGCGGCATACGTGTGACACACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437
480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db TCCCGCCTCTCACCCAAACTACTTGTTCCTTCCTCCCGCGGCACCAAGCAACATGGCC 1497
500 TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db TACGGACCTTCACTTCTCGGGGGCGGCTGATGATCCCTAATACGGGGATCAGCCCTC 1557
520 LeuIleProAspAlaIleProArgGlyLysIleTyrGluIleTyrIleThrLeuHis 539
Db CTCATACCCCGGATGCCATCCCGGAGGAAGATCTACGAGATCTACCTCACACTGCAC 1617
540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db AAGCCAGAAAGCTGAGGTGCCCCCTAGCTGCTGTGACACCCCTGTGATGCTCCAGTCGT 1677
560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db AGCTGTGGGCCCCCAGGAGTCTGCTCACCCGGCAGTATCTCTTGCAATGGACCACTGT 1737
580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db GGAGAGCCCGCCCTGACAGCTGGAGTCTGCGCCTCAAAAGACAGTCTCTGCGAGGGCAGT 1797
600 TrpGluGlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGln 619
Db TGGGAG---GATGCTGACCTTGGTGGAGGAGTCACTTCCCACTCTACTACTGCCAG 1854
620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db CTGGAGGCGGGCCCTGCTATGCTTTCAGGAGAGCTGGGCGCTTGGCCCTGGTAGGA 1914
640 GluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
Db CAGGCGCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCTTCTGCTTCTCCCTGCC 1974
660 CysThrSerLeuTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db TGTACGTCCTTGTAGTACAACATCGAGTGTACTGCTTACAGCACCCACGACGCTCTC 2034
680 LysGluValValGlnLeuGlnLysGlnLeuGlyGlyGlnLeuIleGlnProArgVal 699

2035 AAGGAGTGTGTCAGCTGGAGAGCAGCTAGGTGGACAGCTATCCAGAGCCTCCGCTC 2094
700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db CTGCACTTCAAAAGACAGATTACCACACCTTACGTCTCTCCATCCACGCGTGGCCAGCTCC 2154
720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db CTGTGGAAGAGCAAGCTACTTGTCACTACCAAGGAGATCCCTTTTACCACATCTGGAAC 2214
740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db GGCACCCAGCATCTGCACTTGCACCTTCACTTGGAGGAGATCCACCCAGCCAGCACGAG 2274
760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGAGATGGGCAGAGCTTCAACATC 2334
780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db AACTTCAACATCACTAAGGACACAAAGTTTGTCTGAATTCTTGGCTCTGGAGAGTGAAGG 2394
800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db GGGGTCCACGCTGTGGGGCCCGAGTGCCTTCAAGATCCCTTCTCAATCGGCAAAAG 2454
820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db ATCATCGCCAGTCTGGACCCACCTTGCAGCCGGGGCGGAGCTGGAGAACTCTAGCCAG 2514
840 LysLeuHisLeuAspSerHisLeuSerPheAlaSerLysProSerProThrAlaMet 859
Db AAACCTTCACTGGACAGCATCTTAGCTTCTTGGCTTCCAGCCAGCCCTACAGCCATG 2574
860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAla 879
Db ATCTTCAACCTATGGAGGACGAGCACTTCCCAAGCGGAACCTCGGCAGCTGGCAGCA 2634
880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db GCTGTGGCGGAGTGGCCAAACAGATGCTGSCCTTCTTACGCTGTGGAGGCCGAGTGT 2694

RESULT 6
US-10-256-702-1
; Sequence 1, Application US/10256702
; Publication No. US20030059859A1
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/256,702
; FILING DATE: 27-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261

;
; FILLING DATE: 20-Aug-2001
; APPLICATION NUMBER: 08/808,982
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3014 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-256-702-1

Alignment Scores:

Pred. No.: 0 Length: 3014
Score: 4545.50 Matches: 852
Percent Similarity: 96.78% Conservative: 19
Best Local Similarity: 94.67% Mismatches: 26
Query Match: 94.96% Indels: 3
DB: 15 Gaps: 3

US-09-970-944-2 (1-899) x US-10-256-702-1 (1-3014)

QY	1	MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyLeuValLeuAlaLeuAlaLeuTrpLeu	20
DB	1	ATGGCCGCTCGGGCCCGGCTGTGGCCAGTGTCTCTGGGCATAGTCTCTCGCCGCTGGCT	60
QY	21	ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro	40
DB	61	CGTGGTTCGGTGCCCGACAGAGTGCACGGTGGCCCAATCCAGTGCCTGGTCCCAACCCC	120
QY	41	AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValIysAsnLysPro	60
DB	121	GACCTGTCGCCACATTCCTGTGTAGAGCCTGTGAGAGCTGTACATTTCAAGAACCAAGCCG	180
QY	61	ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu	80
DB	181	GTGTTGTGTGTGCAAGCTGTGCTGTCCACCATCTTCTTCAAGTGCATATGGGAA	240
QY	81	TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro	100
DB	241	TGGGTCCGCGAGTTCGATCACGTAAATTGAACGACGACACGACGAGCGGATTGCCA	300
QY	101	ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu	120
DB	301	ACCATGGAGTCCGATATCAACGTATTCAGGACAGCAGGTAGAGAAAGTGTGGGCTGGAG	360
QY	121	GluTyrTrpCysGlnCysValAlaLeuTrpSerSerGlyThrThrLysSerGlnLysAla	140
DB	361	GAATACTGTGCGAGTGTGTGGATGAGCTCTCGGGTACCCCAAAAGTCAGAGGCC	420
QY	141	TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal	160
DB	421	TACATCCGATTGCCATTTCGCAAGAACTTTGAGCAGGAGCACTGGGCCAAGGAGTG	480
QY	161	SerLeuGlnGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu	180
DB	481	TCATCGAGCAGAGCATTGTACTTCTGTCGCCCCCCCCAGAGAAATCCCCCAGCTGAG	540
QY	181	ValGlnTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle	200
DB	541	GTGAGTGGCTTCGAATGAGGACCTCTGTGGACCCCTCCCTCGATCCCAATGTGTACATC	600
QY	201	ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr	220
DB	601	ACGCGGAGCACAGCCTAGTGTGTGCTGACGGCCCGCTGTGGCCGACACGCGCAACTACCC	660

QY	221	CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr	240
DB	661	TGTGTGGCCAAAGAACATCGTAGCCCGTCCCGAAGCAGCTCTGCAGCGGTCAATTGTTAT	720
QY	241	ValAsnGlyGlyTrpSerThrTrpGluTrpSerValCysSerAlaSerCysGlyArg	260
DB	721	GTGAACGGTGGTGTTCGACGTGGACTGAGTGGTCCGTCTGCAGCCGACAGCTGTGGCGT	780
QY	261	GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe	280
DB	781	GGCTGGCAGAAACGAGCGGAGCTGCACCAACCGGCACCTCTCAACCGGGCGCGCTTC	840
QY	281	CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly	300
DB	841	TGTGAGGGGCAGANTGTC---CAGAAAAACAGCTCGCCACCTCTGTGCCAGTGGATGG	897
QY	301	SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer	320
DB	898	AGCTGGAGTTCGTGGAGTAAGTGTTCAGCTCTGGGCTTGACTGCACCCACTGGCGGAGC	957
QY	321	ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu	340
DB	958	CGCGAGTGTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTGGGGTGTGACCTG	1017
QY	341	AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal	360
DB	1018	GACACCGCACTGTACCACTGACCTCTGCTGCACACCGCTTTGTCGCCCGGAGGACGT	1077
QY	361	AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu	380
DB	1078	GTCTCTACATCGGCTTGTGCTGTGGCTGTGTGCTCTTCTTGTGTGTGGCCCTT	1137
QY	381	IleLeuValTyrCysArgLysGlyGluLeuAspSerAspValAlaAspSerSerIle	400
DB	1138	GGCACTATTACTGTGCAAGAAAGGAGGCTGGACTCCGATGTGGCCGACTCGTCCATC	1197
QY	401	LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu	420
DB	1198	CTACCTCGGGCTTCAGCTGTGAGCATCAAGCCCAAGCAGAACACACCCCACTTG	1257
QY	421	LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThrThr	439
DB	1258	CTCACCATCCAGCCAGACCTCAGCACCACTACCACTACCACTACCACTACCACTAC	1317
QY	440	ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu	459
DB	1318	AGGCAGGATGGACCCAGCCCAAGTTCAGCTCTCTAATGGTCACTGTCTCAGCCACTG	1377
QY	460	GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal	479
DB	1378	GGAGTGGCGCCCATACGTTGCACACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC	1437
QY	480	SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr	499
DB	1438	TCCCGCTCTCCACCAAACTACITTCGTTCCCTGCCCGCGGACCAACCAATGGCC	1497
QY	500	TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu	519
DB	1498	TACGGGACCTTCAACTTCCTCGGGGCGCGCTGATGATCCCTAATAACGGGGATCAGCCTC	1557
QY	520	LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis	539
DB	1558	CTCATACCCCGGATGCCATCCCCGAGGAAGAATCTACAGATCTACCTCACCTGCAC	1617
QY	540	LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal	559
DB	1618	AAGCCAGAGACGCTGAGGTGGCCCTAGCTGGCTGTGCAGACCTCTGTGAGTCCAGTCGT	1677
QY	560	SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys	579
DB	1678	AGCTGTGGGCCCCCAGGAGTCTGTCTCACCCGCGCAGTCTATCTTCAATGGACCACTGT	1737

```
QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db 1738 GGAGAGCCAGCCGCTGACAGCTGGAGTCTGGCTCAAAAAGCAGTCTCGGAGGGCAGT 1797
QY 600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTrpCysGln 619
Db 1798 TGGGAG---GATGTGCTGCACCTTGGTGAGAGTACCTTCCACCTCTACTACTGCCAG 1854
QY 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db 1855 CTGGAGCCGGCCGCTGCTATGCTTTCACGAGCAGCTGGCCGCTTGGCCCTGGTAGGA 1914
QY 640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla 659
Db 1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCTCAGGCTCCTCTGTTTGTCTCCGTGGCC 1974
QY 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db 1975 TGTACGTCCCTTGAGTACAACTCCAGTGTACTGCTTACGACACCCACGACGCTCTC 2034
QY 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
Db 2035 AAGGAGGTGGTGCAGCTGGAGAGCAGCTAGTGGACAGTGTATCCAGGAGCCTCGCGTC 2094
QY 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db 2095 CTGCACCTTCAAAAGACAGTTACCACAACTACGCTCTCTCCATCCAGCAGCTGCCAGCTCC 2154
QY 720 LeuTrpLysSerLysLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db 2155 CTGTGGAAGAGCAAGCTACTTGTACGTACCAAGGAGATCCCTTTTACCACATCTGSAAC 2214
QY 740 GlyThrGlnArgTyrIleLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db 2215 GGCACCCAGCAGTATCTGCACCTTGACCTTCACTGGAGCGCATCAACCCAGCAGCACG 2274
QY 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db 2275 GACCTGGCCCTGCAAGGTGGGTGGTGGCAGGTGGAGGAGATGGGCAGAGCTTCAACATC 2334
QY 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuAlaLeuGluSerGluAla 799
Db 2335 AACTTCAACATCACAAGGACACAAGTGTGTGAATTTGTGGCTCTGGAGAGTGAAGGG 2394
QY 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db 2395 GGGGTCCAGCCCTGGTGGGCCCGCAGTGCCTTCAAGATCCCTTCTCATTCGGCAAAAG 2454
QY 820 IleLeuSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db 2455 ATCATCGCCAGTCTGGACCCACCTCGACCGGGGCGCGACTGGAGAACTTAGCCCGAG 2514
QY 840 LysLeuHisLeuAspSerHisLeuSerPheAlaSerLysProSerProThrAlaMet 859
Db 2515 AAACCTTCACTGGACAGCATCTTAGCTTCTTTGTCTTCAACCCAGCCGCTACAGCCATG 2574
QY 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAla 879
Db 2575 ATCCTCAACCTATGGAGGCGCAGGCACTTCCCAACGGCAACCTCGGCAGCTGGCAGCA 2634
QY 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGlyCys 899
Db 2635 GCTGTGGCCGAGTGGGCCAACAGATGCTGSCCTTTCACGGTCTCGGAGGCCAGGTGT 2694
RESULT 7
US-10-311-623-13
; Sequence 13, Application US/10311623
; Publication No. US2004002344A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.
; APPLICANT: KALLIUCK, Deborah A.; TRIBOULEY, Catherine M.
; APPLICANT: YUE, Henry; NGUYEN, Darniel B.
; APPLICANT: TANG, Y. Tom; LAL, Preeti G.
```

```
; APPLICANT: POLICKY, Jennifer L.; AZIMZAI, Valda
; APPLICANT: LU, Dyung Aina M.; GRAUL, Richard C.
; APPLICANT: YAO, Monique G.; BURFORD, Neil
; APPLICANT: HAFALIA, April J. A.; BAUGHN, Mariah R.
; APPLICANT: BANDMAN, Olga; ARVIZU, Chandra S.
; APPLICANT: YANG, Junming; XU, Yuming
; APPLICANT: GANDHI, Ameena R.; WARREN, Bridget A.
; APPLICANT: DING, Li; SANJANWALA, Madhusudan M.
; APPLICANT: DOUGGAN, Brendan M.; LU, Yan
; TITLE OF INVENTION: RECEPTORS
; FILE REFERENCE: PF-0793 USN
; CURRENT APPLICATION NUMBER: US/10/311,623
; PRIOR FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US 01/19942
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/214,027
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/228,045
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/255,104
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 3580
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US2004002344A1 6052371CB1
US-10-311-623-13

Alignment Scores:
Pred. No.: 0 Length: 3580
Score: 4379.50 Matches: 838
Percent Similarity: 93.22% Conservative: 1
Best Local Similarity: 93.11% Mismatches: 2
Query Match: 91.43% Indels: 59
DB: 17 Gaps: 3

US-09-970-944-2 (1-899) x US-10-311-623-13 (1-3580)
QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaIleTrpLeu 20
Db 4 ATGGCCGTCGCGCCCGGCTGTGGCCAGCGTCTGGGGCATAGTCTCGCCCGCTTGGCTC 63
QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db 64 CGCGGCTCGGGTGGCCAGCAGAGTGCACCGTGGCCACCCAGTGCCTGGTGCCACCCG 123
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db 124 GACCTGCTTCCCACTTCTCTGTGGAGCCGAGGATGTATCATCTCAAGAACCAAGCCA 183
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db 184 GTGCTGTGTGTGCAAGCGCGTGGCCCGCCAGCATCTTCTTCAAGTGAACCGGGAG 243
QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerGlyGluPro 100
Db 244 TGGGTGGCCAGGTGGACCACTGATCGAGCGCAGCAGACGCGGAGCAGTGGGTGGCC 303
QY 101 ThrMetGluValArgIleAsnValSerArgGlnValGluLysValPheGlyLeuGlu 120
Db 304 ACCATGGAGGTCCGCATTAATGTCTCAAGGCGCAGGTCGAGAAGTGTTCGGGCTGGAG 363
QY 121 GluTyrTrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAla 140
Db 364 GAATACTGGTCCAGTGGTGGCATGGAGCTCTCGGGCACCAACCAAGAGTCAGAAGGCC 423
QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db 424 TACATCCGCATAGCTTATTTGCGCAAGNACTTTCAGCAGGAGGCCCTGCGCCAGAGGTG 483
```

QY	161	SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu	180	QY	520	LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis	539
DB	484	TCCTTGGAGCAGGCATCGTGCTGCCCTGCCGTCCACGGAGGCATCCCTCCAGCCGAG	543	DB	1393	CTCATCCCCCAGATGCCATACCCCGAGGAGAGATCTAGATCTACTCAGCGTGCAC	1452
QY	181	ValGluTyrLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle	200	QY	540	LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal	559
DB	544	GTGGAGTGGCTCCGGAACAGAGACCTGGTGGACCCGTCCTCGACCCCAATATACATC	603	DB	1453	AAGCCGGAAGAGCTGAGGTGCCCTAGCTGGCTGTGCAGACCCTGCTAGTCCCATCGTT	1512
QY	201	ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr	220	QY	560	SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys	579
DB	604	ACGGGGAGCACAGCTGTGTGTGGACAGGCCCGCTGTGTGCACGGCCAACTACACC	663	DB	1513	AGCTGTGAGCCCTTGGGTCCTGTCTACCCGGCCAGTCATCTGGCTATGACCACTGT	1572
QY	221	CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr	240	QY	580	GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer	599
DB	664	TGCGTGGCCACAGAACATCGTGACCGTCGCCCGCAGCGCTCGCTGCTCATCGCTAC	723	DB	1573	GGGGAGCCAGCCCTTGACAGCTGGAGCCTGCCCTCAAAAAGACGCTGTCGAGGGCAGC	1632
QY	241	ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg	260	QY	600	TrpGluGlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGln	619
DB	723	-----	723	DB	1633	TGGGAG--GATGTGCTGCACCTGGCGGAGGAGGGCCCTCCGACCTCTACTACTGCCAG	1689
QY	261	GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyAlaPhe	280	QY	620	LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly	639
DB	723	-----	723	DB	1690	CTGGAGGCCAGTGCCTGTCTTACCCGAGCAGCTGGCGCGCTTTGCCCTGGTGGGA	1749
QY	281	CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly	300	QY	640	GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla	659
DB	724	-----	732	DB	1750	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCTCAAGCTGCTTCTGTTGCGCGGTGGCC	1809
QY	301	SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer	320	QY	660	CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu	679
DB	733	AGCTGAGGCCGTGGAGCAAGTGGTGGCTGTGGGCTGGACTGCACCCACTGGCGGAGC	792	DB	1810	TGCACCTCCCTCGAGTACAAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCATC	1869
QY	321	ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu	340	QY	680	LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal	699
DB	793	CGTGTGCTCTGCACCCAGCACCCCGCAACGGAGGGGAGGTGCCAGGGCACTGACCTG	852	DB	1870	AAGGAGGTGGTCAGCTGGAGAGCAGCTGGGGGGACAGCTGATCCAGAGGCCACGGGTC	1929
QY	341	AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal	360	QY	700	LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer	719
DB	853	GACACCGCAACTGTACCACTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTG	912	DB	1930	CTGCACCTTCAGGACAGTTTACCACACCTTGGCCCTATCCATCCACGATGTGCCCAGCTCC	1989
QY	361	AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu	380	QY	720	LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn	739
DB	913	GCCTCTATGTGGGCTCATCGCGGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTC	972	DB	1990	CTGTGGAAGAGTAAGCTCTTGTACGTACCAGGAGATCCCTTTTATCACATCTGGAAT	2049
QY	381	IleLeuValTyrCysArgLysLysGlyLeuAspSerAspValAlaAspSerSerIle	400	QY	740	GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProThrThrSer	759
DB	973	ATCCTCGTTTATTCGCGAAGAGAGGGGCTGGACTCAGATCTGCTGACTCGTCCATT	1032	DB	2050	GGCACGACGGTACTTGCACCTGCACCTTACCCCTGGAGCGGTGACGCCCCAGCACTAGT	2109
QY	401	LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu	420	QY	760	AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle	779
DB	1033	CTCACCTCAGGCTTCAGCCCGCTCAGCATCAAGCCAGCAAGCAGACACACCCCACTG	1092	DB	2110	GACCTGGCTTCNAGCTGTGGGTGTGGAGGTGGAGGGCGAGCGGCGAGCTTCAGCATC	2169
QY	421	LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThrThrThr	439	QY	780	AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla	799
DB	1093	CTCACCATCCAGCGGACCTCAGACCAACACACACACCTACACGGGCGAGTCTCTGTCCC	1152	DB	2170	AACTTCAACATCACCAAGGACACACAGGTTTGTGTAGCTGTGTGTCTGTGGAGGTGAAGCG	2229
QY	440	ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu	459	QY	800	GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys	819
DB	1153	CGGCAGGATGGGCCAGCCCCCAAGTTCCAGCTCACCAATGGGCACCTGTCTCAGCCCCCTG	1212	DB	2230	GGGGTCCCAGCCCTGGTGGGCCCCAGTGTCTTCAGATCCCTTCTCATTCGCGCAGAAG	2289
QY	460	GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluPheVal	479	QY	820	IleIleSerSerLeuAspProProCysArgArgGlyValaAspTrpArgThrLeuAlaGln	839
DB	1213	GGTGGGGCGCCACACACTGGACACACAGCTCTCCACACCTCTGAGGCCGAGGAGTTGCTC	1272	DB	2290	ATAATTTCCAGCTTGGACCCACCCCTGTAGCGGGGTGCCGACTGGCGGACTCTGGCCAG	2349
QY	480	SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr	499	QY	840	LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet	859
DB	1273	TCCGGCTCTCCACCCGAACTACTTCGGTCCCTGCCCGAGGCACCAACATGACC	1332	DB	2350	AAACTCCACTGACAGCCATCTCAGCTTCTTTGCTCCAAAGCCCGACCCACAGCCATG	2409
QY	500	TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu	519	QY	860	IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla	879
DB	1333	TATGGGACCTTCAACTTCTTCGGGGCGGCTGTATGATCCCTAATACAGGAATCAGCCTC	1392	DB	2410	ATCCTCAACTGTGGGAGGGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2469
DB				QY	880	AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys	899

Db 2470 GCAGTGGCTGGAGTGGCCAGCAGCTGGCTCTTACAGTGTGGAGGCTGAGTGC 2529
|||||
RESULT 8
US-09-933-261-2
; Sequence 2, Application US/09933261
; Publication No. US20030040046A1
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261
; Filing DATE: 20-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/808,982
; Filing DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1787 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-933-261-2
Alignment Scores:
Pred. No.: 5,35e-274 Length: 1787
Score: 2694.00 Matches: 543
Percent Similarity: 97.14% Conservative: 1
Best Local Similarity: 96.96% Mismatches: 11
Query Match: 56.28% Indels: 11
DB: 10 Gaps: 2
US-09-970-944-2 (1-899) x US-09-933-261-2 (1-1787)
Qy 344 AsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363
Db 3 AACTGTACAGTACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCTAT 61
Qy 364 ValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeuVal 383
Db 62 GTGGGCTCATCGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
Qy 384 TyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSer 403
Db 122 TATTGCCGGAAGAGGAGGGGTGACTCAGATGGCTGACTCGTCCATTCTCACCTCA 181

Qy 404 GlyPheGlnProValSerIle-LysProSerLysAlaAspAsnProHisLeuLeuThrI 423
Db 182 GGCTTCCAGCCCGTACGATCTAAGCCCAAGCAGCAACCCCATCTGCTCCACAT 241
Qy 423 eGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysProArgGlnAs 442
Db 242 CCAGCGGACCTCAGCACCAACACCACTACAGGGGAGTCTCTGTCTCCCGGAGGA 301
Qy 442 pGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGly 462
Db 302 TGGGGCCAGCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCG 361
Qy 462 VArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSerArgLe 482
Db 362 CGGCCACACACTGCACCACTCTCCACCTCTGAGGGCCGAGGAGTTCGTCTCCCGCT 421
Qy 482 uSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyrGlyTh 502
Db 422 CTCACCCAGAACTACTTCGGCTCCCTGCCCGAGGCACAGCAACATGACCTATGGAC 481
Qy 502 rPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePr 522
Db 482 CTTCAACTTCTCGGGGGCCGCTGATGATCCTAATACAGGAATCAGCTCTCATGCC 541
Qy 522 oProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProG 542
Db 542 CCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAGCTGCACAAGCCGA 601
Qy 542 uAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGl 562
Db 602 AGACCTGAGTTGCCCTAGCTGGCTGTGAGACCTCTGAGCTCCATCGTTAGCTGTGG 661
Qy 562 yProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyLupr 582
Db 662 ACCCTCGCGTCTGCTCACCGGCCAGTATCTGCTGCTATGGACCACTGTGGGGAGCC 721
Qy 582 sSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGl 602
Db 722 CAGCCCTCAGACGTGGAGCTGGCCCTCAAAAGCAGTCTGTCGAGGGAGC-TGGAG-- 778
Qy 602 nAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAl 622
Db 779 -GATGT-CTGCACCTGGGGAGGAGGCCCTCCACCTCTACTACTCGCAGCTGGAGCC 836
Qy 622 aserAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlale 642
Db 837 CAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGGCGCTTTGCCCTGGTGGGAGGCCCT 896
Qy 642 uSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAlaCysThrSe 662
Db 897 CAGCGTGGCTCGGCCAAAGCGCTCAAGCTGCTTCTGTTGCGCGGTGGCTGCACCTC 956
Qy 662 rLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluVa 682
Db 957 CCTCGAGTACAACTCCGGTCTACTGCTGATGACACCCACCATGACCTCAAGAGAGT 1016
Qy 682 lValGlnLeuGluLysGlnLeuGlyGlnLeuIleGlnGluProArgValLeuHisPh 702
Db 1017 GGTGCAGCTCGAAGAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCACCTC 1076
Qy 702 eLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLy 722
Db 1077 -AAGGACAGTTACCAACACCTGCC-CTATC-ATCCACGATGTGCCACCTCCCTGTGGA 1133
Qy 722 sSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGl 742
Db 1134 GAGTAAGCTCTTGTGCTACCTACAGGAGATCCCCCTTTTATCACATCTGGAATGCACG 1193
Qy 742 nArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAl 762
Db 1194 GCGGTACTTGCACTGCACCTTCACCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGC 1253
Qy 762 aCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAs 782

Db 1254 CTGCAAGCTGTGGGTGTGGCAGTGGAGGGCGACGGGCGAGAGCTTCAGCATCAACTTCAA 1313
QY 782 nleThrLysAspThrArgPheAlaGluLeuAlaLeuGluSerGluAlaGlyValPr 802
Db 1314 CATCAACAGGACACAAAGGTTTGCTGAGCTGTGGCTCTGGAGAGTGGAAGCGGGGTCCC 1373
QY 802 oAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleSe 822
Db 1374 AGCCCTGTGGTGGCGCCAGTGCCTTCAAGATGCCCTTCTCTTGGCAGAGATAATTTC 1433
QY 822 rSerLeuAspProProCysArgGlyAlaAlaAspThrArgThrLeuAlaGlnLysLeuHi 842
Db 1434 CAGCCTGGAGCCACCCCTGTAGGGGGGTGCCGACTGTGGCGAGCTCTGCCCCAGAACTCCA 1493
QY 842 sLeuAspSerHisLeuSerPheAlaSerLysProSerProThrAlaMetIleLeuAs 862
Db 1494 CTTGGACAGCCATCTCAGCTTCTTTGGCTTCCAAAGCCACGCCCCACAGCCATGATCCTCAA 1553
QY 862 nLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAl 882
Db 1554 CCGTGGGAGGGCGGCGACTTCCCAACGGCAACCTCAGCCAGCTGGCTGGCAGAGTGGC 1613
QY 882 acLysLeuGlyGlnProAspAlaGlyLeu-PheThrVal-SerGluAlaGluCys 899
Db 1614 TGGGACTGGCCAGCAGGACGGTGGCTTCTTTACAGTGTTCCGAGGCTGAGTGC 1667

RESULT 9

US-10-256-702-2
; Sequence 2, Application US/10256702
; Publication No. US20030059859A1
; GENERAL INFORMATION:
; APPLICANT: Tessier-Javigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/256,702
; FILING DATE: 27-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261
; FILING DATE: 20-Aug-2001
; APPLICATION NUMBER: 08/808,982
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1787 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

; MOLECULE TYPE: CDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-256-702-2
Alignment Scores: Length: 1787
Pred. No.: 5,35e-274 Matches: 543
Score: 2694.00 Conservative: 1
Percent Similarity: 97.14% Mismatches: 11
Best local Similarity: 96.96% Indels: 11
Query Match: 56.28% Gaps: 2
DB: 15
US-09-970-944-2 (1-899) x US-10-256-702-2 (1-1787)
QY 344 AnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363
Db 3 AACTGTACCAAGTACCTCTG-GTACACACTGTCTTGGGCCCTGGAGACGTGGCCCTCTAT 61
QY 364 ValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeuVal 383
Db 62 GTGGGCTCATCGCGTGGCGCTTGGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
QY 384 TyrCysArgLysGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSer 403
Db 122 TATTGCCGGAAGAGAGGGGCTGGACTCAGATGTGGTGACTCGTCCATTCTCACTCA 181
QY 404 GlyPheGlnProValSerIle-LysProSerLysAlaAspAsnProHisLeuThrI 423
Db 182 GGCCTTCCAGCCGCTCAGCATCTAAGCCAGACAGACAAACCCCATCTGCTCACTAT 241
QY 423 eGlnProAspLeuSer---ThrThrThrThrGlnGlySerLeuCysProArgGlnAs 442
Db 242 CCAGCGGACCTCAGACACACACACACACCTACAGGGCAGTCTCTGTCCCGGCAGGA 301
QY 442 pGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGly 462
Db 302 TGGGGCCAGCCCAAGTTCAGCTCACCATAATGGGCACCTGCTCAGGCCCTGGGTGGCGG 361
QY 462 yArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSerArgLe 482
Db 362 CGGCCACACACTGCACACAGCTCTCCACCTCTGAGGGCGAGGAGTGTGTCTCCCGCCT 421
QY 482 uSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyrGlyTh 502
Db 422 CTCACCCAGAACTACTTCCGCTCCCTGCCCGGAGCACCAACATGACCTATGGGAC 481
QY 502 rPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePr 522
Db 482 CTTCAACTTCTCGGGGGCGGCTGATGATCCCTAATACAGGAATCAGCCCTCCTCATCCC 541
QY 522 oProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGl 542
Db 542 CCAGATGCCATACCCCGGAGGAAGATCTATGATCTACCTCAGCTGTCAGTCCCATCGTAGCTGG 601
QY 542 uAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGl 562
Db 602 AGACGTGAGTTGCCCTTAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
QY 562 yProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluPr 582
Db 662 ACCCCCTGGCGTCTGCTCACCGGGCAGTCACTCTGGCTATGGACACTGTGGGAGGCC 721
QY 582 oSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGl 602
Db 722 CAGCCCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTGTGTGGAGGAGC-TGGAG-- 778
QY 602 nAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAl 622
Db 779 -GATGT-CTGCACCTGGCGGAGGAGGCGCCCTCCACCTCTACTACTGTCAGCTGGAGGC 836
QY 622 aSerAlaCysTyrValPheThrGlnLeuLeuGlyArgPheAlaLeuValGlyGluAlaLe 642
Db 837 CAGTGGCTGTCTCTTTCACCCAGCAGCTGGGCGCGCTTTGGCCCTGTGTGGAGAGGCCCT 896


```
Db 484 GCGGGACACCAAGAGTCCGCGAGCTAGCTCCGCATCGCCTGTCTGCGCAAGACTTC 543
Qy 152 GluGlnGluProLeuAlaLysGluValSerLeuGluGlnGlyLeValLeuProCysArg 171
Db 544 GATCAGGAGCCTCTGGCGCAAGAGGTGCCCTTGACCATGAGGTTCCTCTGCAGTGGCG 603
Qy 172 ProProGlnGlyLeProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAsp 191
Db 604 CCGCCCGAGGGGTGCTCTGCGCGAGGTGGATGGCTCAAGAATGAGGATGTCATCGAC 663
Qy 192 ProSerLeuAspProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAla 211
Db 664 CCCACCCAGCAGCACCACTTCCTGCTCACCATCGACCACTCATCATCGCCAGGCC 723
Qy 212 ArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArg 231
Db 724 CGCCTGTGCGACACTGCCAACTATACCTGCTGGCCCAAGAACATCGTGGCCAAACGCCG 783
Qy 232 SerAlaSerAlaAlaValIleValTyrValAsnGlyTrpSerThrTrpThrGluTrp 251
Db 784 AGCACCACTGCCACCGCTCATGCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGG 843
Qy 252 SerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsn 271
Db 844 TCACCTGTCTCAACCGCTGTGGCCGAGGTGGCAGAGCGCACCCGGACCTGCACCAAC 903
Qy 272 ProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrVal 291
Db 904 CCGCTGCCACTCAACGGAGGGGCTTCTGCGAGGGCGAG---GCATTCCAGAAAGACCGCC 960
Qy 292 SerSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysLysTrpSerAlaCys 311
Db 961 TGCACCACTATGCCCGATGCGATGGGGCTGGACGAGTGGAGAGTGGTCAAGCTGC 1020
Qy 312 GlyLeuAspCysThrHisTrpArgSerArgGlyCysSerAspProAlaProArgAsnGly 331
Db 1021 AGCACTGAGTGTGCCCACTGGCGTAGCGCGAGTGCATGGCGCCCCCACCCAGACGGA 1080
Qy 332 GlyGluGlyCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysVal 351
Db 1081 GGCGGTGAGTGCAGCGGAGCGCTGCTCGACTTAAGAACTGCACAGATGGGCTGTGCATG 1140
Qy 352 HisSer-----AlaSerGlyProGluAspValAlaLeuTyr 363
Db 1141 CAAAGTGAGCTGTCCCGCAGTCTGGAGGCTCAGGG-----GATGGCGGCTGTAT 1194
Qy 364 ValGlyLeu---IleAlaValAlaValCysLeuValLeuLeuLeuValLeuLeu 382
Db 1195 GCGGGGCTGCTGTGGCCATCTTCGTGTCGTGGCAATCTCATGGCGGTGGGGTGGTG 1254
Qy 383 ValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSer---IleLeu 401
Db 1255 GTGTACCGCGCAACTGCCGTGATCTTCGACACAGACATCACTGACTCATCTGTGCGCCCTG 1314
Qy 402 ThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu 421
Db 1315 ACTGGTGGTTCCACCCCGTCAACTTAAGACCGCAAGGCCAGTAACCCGCGAGCTCTTA 1374
Qy 422 -----ThrIleGlnProAspLeuSerThrThrThr-----ThrTyrGlnGlySerLeuCys 438
Db 1375 CACCCCTCTGTGCTCTGACTGACAGCGCGCGGACATCTACCGCGGACCGGTAT 1434
Qy 439 ProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerPro 458
Db 1435 GCCTTGAGGACTTCC---ACCGCAAAATCCCATGACCAACTCTCTCTGTCTGGACCCC 1491
Qy 459 LeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThr----- 472
Db 1492 TTACCCAGCCTTAAGTCAAGGTCTAGCTCCAGCACCCAGGCTCTGGCGAGGCGCTG 1551
Qy 473 SerGluAlaGluGluPheValSerArgLeuSerThrGlnAsnTyr----- 487
Db 1552 GCATGGGGTGTACCTGCTGGGGGTCTTGGCGGCTGGCACATACCTAGCGATTGTCGCC 1611
```

```
Qy 488 -----PheArgSer----- 490
Db 1612 CCGGACACCCACTTCTGCACCTGGCAGCGCAGCCTCGGTTCGCCAGCAGCTCTTGGGCG 1671
Qy 491 LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeu 510
Db 1672 CTGCCCCGAGACCCAGGAGCGCTCAGCGGCACCTTGGCTGCTGGGTGGAGGCTC 1731
Qy 511 MetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLys 530
Db 1732 AGCATCCCCCGCACAGTGTGCTGCTGGTGCCCATGAGCCATTCGCCAGGCAAG 1791
Qy 531 IleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla--- 549
Db 1792 TTCTACGAGATGTATCTACTATCAACAAGCGCAAGAAATACC---GTGCCGCTTTCAGAA 1848
Qy 550 GlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyValLeuLeuThr 569
Db 1849 GGGACCCACACAGTATTGAGCCCTCGGTGACCTGTGGACCCACAGGCTCTCTGTGTGC 1908
Qy 570 ArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerThrSerLeu 589
Db 1909 CGCCCCGTATCTCACCATGCCCACTGTGCCGAAGTCAGTCCCGCTGACTGGATCTTT 1968
Qy 590 ArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGlu 609
Db 1969 CAGCTCAAGACCCAGGCCCCACAGGCCCCACTGGAGCAGAGGTGTGACCTTGGATGAG 2028
Qy 610 GluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThr 629
Db 2029 GAGACCTCAACACACACCTTCTACTGCGAGCTGGAGCCCGGCTGTACATCTCTGCTG 2088
Qy 630 GluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaLysArg 649
Db 2089 GACCAGCTGGCACCTACGTGTTACGGGCGAGTCTCTATTCCGCTCAGCAGTCAAGCGG 2148
Qy 650 LeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgVal 669
Db 2149 CTCACAGCTGGCTCTTCGCCCGCCCTCTGCACCTCTCTGGAGTACAGCTTCCGGGTC 2208
Qy 670 TyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeu 689
Db 2209 TACTGCTGGAGGACAGCGCTGTAGCAGTGAAGAGGTGTGTGAGCTGGAGCGGACTCTG 2268
Qy 690 GlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709
Db 2269 GCGGATACTTGTGTGGAGGCGCAACCGCTAATGTTCAAGGACAGTTTACCACCAACCTG 2328
Qy 710 ArgLeuSerIleHisAspValProSerSerLeuTyrLysSerLysLeuLeuValSerTyr 729
Db 2329 CGCCTCTCCTCCATGACCTCCCTCCATGCCATGCGCATGGAGGAGCAAGTGTGGCCAAATAC 2388
Qy 730 GlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPhe 749
Db 2389 CAGGAGATCCCCCTTCTATCACAATTTGGAGTGGCAGCAGAAAGGCCCTCCACTGCATCTT 2448
Qy 750 ThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGln 769
Db 2449 ACCCTGGAGGCGACAGTGTGGCTCCACAGAGTCACTGTCAAGATCTCGTGGCGGCAA 2508
Qy 770 ValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPhe 789
Db 2509 GTGGAGGGGAGGCGCCAGATATTCAGCTGCATACCACTCTGSCA---GAGACACCTGCT 2565
Qy 790 AlaGluLeuAlaLeuGluSerGluAlaGly-----valProAlaLeuValGlyPro 807
Db 2566 GGCTCCTCGACACTCTCTGCTGTGCCCTGGCAGCAGCTGTCAACCCAGCTGGGACCT 2625
Qy 808 SerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProPro 827
Db 2626 TATGCTTCAAGATCCCACTGTCCATCCGCGCAGAGATATGCAACGCTAGATGTCGCC 2685
```

```

QY      828 CysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeu 847
Db      2686 AACTCAGGGGGCAATGACTGGCGGATGTAGCACAGAGCTCTCTATGGACCGGTACCTG 2745
QY      848 SerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArg 867
Db      2746 AATTACTTTGCCACCAAGAGCGAGCCCGGGTGTGATCTGGACCTCTGGGAAGCTCTG 2805
QY      868 HisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnPro 887
Db      2806 CACGACAGACATGGGACCTCAACAGCCCTGGCGAGTGCCTTGGAGAGATGGCAAGAGT 2865
QY      888 AsPalaglyLeuPheThrValSerGluAlaGluCys 899
Db      2866 GAGATGCTGTGGCTGTGGCCACCGAGGGGACTGC 2901

```

RESULT 12

```

US-10-087-684-1
; Sequence 1, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: MacDougall, John R.
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David J.
; APPLICANT: Grosse, William M.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie, J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zernhusen, Bryan D.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles E.
; APPLICANT: Gangolli, Esha A.
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-214 CIP
; CURRENT APPLICATION NUMBER: US/10/087,684
; PRIOR FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 1
; LENGTH: 2860
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (59)..(2857)
US-10-087-684-1

```

Alignment Scores:

```

pred. No.: 1,72e-251 Length: 2860
Score: 2484.00 Matches: 493
Percent Similarity: 68.30% Conservatives: 149
Best Local Similarity: 52.45% Mismatches: 250
Query Match: 51.89% Indels: 48
DB: 13 Gaps: 18

US-09-970-944-2 (1-899) x US-10-087-684-1 (1-2860)

QY      1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaA1aTrp--- 19
Db      59 ATGGGGGCGCGGAGCGAGCTCGGGCGCGCTGCTGTCGCACTGCTGCTCTGGGAC 118
QY      20 -----LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAsnProValPro 36
Db      119 CCGAGGCTGAGCCAAGCAGGCACTGATTCTGGCAGCGAGGTGCTCCTCTTCCTCCG 178
QY      37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTrIleVal 56
Db      179 TCAGCGCCAGCAGAGCGGCTGCTTCTTCTGAGGAGCCACAGACGCCCTACATITIG 238
QY      57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
Db      239 AAGAACAGAGCTGTGGAGCTCCGCTCGCGCGCTTCCCGCCACACAGATCTACTTCAAG 298
QY      77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
Db      299 TCGAACGGCGAGTGGGTGAGCCAGCAGCACGTCACACAGGAAGGCCCTGGATGAGGCC 358
QY      97 SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysVal 116
Db      359 ACCGGTCTCGGGTGGCGGAGGTGAGATCGAGGTGTCGGCAGCGAGCTCCGCGGAGGAGCTC 418
QY      117 PheGlyLeuGluGluTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLys 136
Db      419 TTTGGCTGGAGGATTACTGTGTCAGTGGCTGGCTGGAGTCCGCGGAGCTCCGCGGAGGAGCTC 478
QY      137 SerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeu 156
Db      479 AGTCGCGGAGCCTACGTCGTCATCGCTACCTCGCAAGAACTTCGATCAGGAGCCTCTG 538
QY      157 AlaLysGluValSerLeuGluGlnGlnIleValLeuProCysArgProGluGlyIle 176
Db      539 GCGAAGAGGAGTCCCTTGACCATGAGGTTCCTCTGAGTGCCTCGCGCGGAGGGGGTG 598
QY      177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196
Db      599 CCTGTGGCGGAGGTGGAATGGCTCAAGATGAGGATGTCTATCGACCCACCCAGGACACC 658
QY      197 AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216
Db      659 AACTTCTCTGCTCACCATGACCAACCTCATCTCCGCGAGCGCGCTGTCGAGACAT 718
QY      217 AlaAsnTyrThrCysValAlaLysAsnIleValAlaArgAtgArgSerAlaSerAlaA1a 236
Db      719 GCCAACTATACCTGCTGGCCAAAGAACTCTGTGGCCAAACGCCGAGCAGCACCCTGCCACC 778
QY      237 ValIleValTyrValAsnGlyGlyTrpSerThrTrpThrTrpGluTrpSerValCysSerAla 256
Db      779 GTCATCGTCTACGTCATGGGGCTGGTCCAGCTGGGCGAGAGTGGTCACTCCCTGCTCCAAC 838
QY      257 SerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276
Db      839 CGCTGTGGCGGAGGTGGCAAGCGCACCCGAGCCTGCACCAACCCCGCTCCACTCAAC 898
QY      277 GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuVal 296
Db      899 GGAGGGGCTTCTGCGAGGGCCAG---GCATTCCAGAAAGACCGCTGCACCACTATCTGCG 955
QY      297 SerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThr 316
Db      956 CCAGTTCGATGGGGCTGGCAGGAGTGGAGCAAGTGGTTCAGCTGCAGCAGCTGAGTGTGCC 1015

```

QY 317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGln 336
Db 1016 CACTGGCGTAGCGCGAGTGCATGGCGCCGCCACCCAGAACGGAGCCGCTGACTCAGC 1075
QY 337 GlyTrpAspLeuAspThrArgAsnCysThrSerAspLeuCysValHis---SerAlaSer 355
Db 1076 GGGACGCTGCTGACTCTAGAACTGCACAGATGGGCTGTGCATGCACTGAGGCGCTCA 1135
QY 356 GlyProGluAspValAlaLeuTyValGlyLeu---IleAlaValAlaValCysLeuVal 374
Db 1136 GGG-----GATCGCGCGTGTATGCGGGCTCGTGTGGCCATCTTCGTGTGTCGGCA 1189
QY 375 LeuLeuLeuLeuValLeuLeuValTyCysArgLysGlyGluGlyLeuAspSerAsp 394
Db 1190 ATCCTCATCGCGTGGGGTGGTGTACCGCGCAACTGCGCTGACTTCGACACAGAC 1249
QY 395 ValAlaAspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSer 413
Db 1250 ATCACTGACTCATCTGCTGCCCTGACTGTGTGTTTCCACCCGCTCAACTTTAAGACGGCA 1309
QY 414 LysAlaAspAsnProHisLeuLeu-----ThrIleGlnProAspLeuSerThrThr 431
Db 1310 AGGCCAGTAACCCGAGCTCTACACCCCTCTGTGCTCTCTGACCTGACGACGCGC 1369
QY 432 ---ThrTyrglnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeu 450
Db 1370 GGCATCTACCGCGGACCGGTGTATGCCCTGCAGACTCC---ACCGACAAATCCCGCATG 1426
QY 451 ThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArgHisThrLeuHisHisSerSer 470
Db 1427 ACCAACTCTCTGTGTGGACCTTACCAGCCTTAAGTCAAGTCTACAGCTCCAGC 1486
QY 471 ProThr-----SerGluAlaGluGluPheValSerArgLeuSerThr 484
Db 1487 ACCACGGGCTCTGGCGAGGCTGGCAGATGGGGCTGACCTGTGGGGTCTTTCGCGCCT 1546
QY 485 GlnAsnTy-----PheArgSer----- 490
Db 1547 GGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGGCGCAGCGCAGC 1606
QY 491 -----LeuProArgGlyThrSerAsnMetThrTyrglyThr 502
Db 1607 CTGGTTCACGAGCTCTTGGGCTGCCCCGAGACCCAGGAGCAGCGCTCAGCGGCACC 1666
QY 503 PheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePro 522
Db 1667 TTTGGCTGCTGGGTGGAGGCTCAGCATCCCGGACACAGGGGTCACTGTGTGGTGGCC 1726
QY 523 ProAspAlaIleProArgGlyLysIleTyrgluIleTyrgluLeuThrLeuHisLysProGlu 542
Db 1727 AATGGAGCCATCCCGAGGCAAGTTCTACGAGATGATCTCATCAACAAGGCGAGAA 1786
QY 543 AspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSerProIleValSerCys 561
Db 1787 AGTACC---CTGCGCGCTTCAGAAAGGACCCACAGACAGTATTGAGCCCTCGGTGACTGT 1843
QY 562 GlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGlu 581
Db 1844 GGACCCACAGGCTCTGTGTGGCGCCCGCTATCTCTCAACATGCCCCCACTGTGGCGAA 1903
QY 582 ProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGlu 601
Db 1904 GTACGTGCCGTGACTGTGATCTTTCAGTCAAGACCCAGGCCCCACCGGCCACTGGGAG 1963
QY 602 GlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyTyCysGlnLeuGlu 621
Db 1964 ---GAGGTGTGACCTCGATGAGGACCCCTGAACACACCTGCTACTGCCAGCTGGAG 2020
QY 622 AlaSerAlaCysTyTrpValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAla 641
Db 2021 CCCAGGGCGCTCACAATCTGCTGGACCACTGCTGGCAGCTACGTGTTCACGGCGAGTCC 2080
QY 642 LeuSerValAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThr 661

Db 2081 TATTCGCGCTCAGACTCAAGCGGCTCCAGCTGGCGCTTTCGCCCCCGCTCTGCACC 2140
QY 662 SerLeuGluTyTrpAsnIleArgValTyCysLeuHisAspThrHisAspAlaLeuLysGlu 681
Db 2141 TCCCTGGAGTACAGCTCGGGTCTACTGCTGGAGACACGCTGTAGCACTGAAGGAG 2200
QY 682 ValValGlnLeuGluLysGlnLeuGlyGlnLeuLeuLeuGlnGluProArgValLeuHis 701
Db 2201 GTGCTGGAGCTGAGCGGACTCTGGCGGATATCTTGTGGAGGAGCGAAACGCTAATG 2260
QY 702 PheLysAspSerTyTrpHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrp 721
Db 2261 TTCAGGACAGTTACACAACTGCGCTCTCCCTCCATGACCTCCCGCATGCCATTGG 2320
QY 722 LysSerLysLeuLeuValSerTyrglnGluIleProPheTyHisIleTrpAsnGlyThr 741
Db 2321 AGGAGCAAGCTGCTGCCCAATATACAGAGAGATCCCTTCTATCACATTTGGAGTGGCAGC 2380
QY 742 GlnArgTyTrpLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeu 761
Db 2381 CAGAAGGCCCTCCACGCACTTTCACCTGGAGAGGCAAGCTTGGCCTCCACAGAGCTC 2440
QY 762 AlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPhe 781
Db 2441 ACTGCAAGATCTGCTGCGGCAAGTGAGAGGGAGGCGGCAGATATTCAGCTGCATACC 2500
QY 782 AsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuSerGluAlaGly--- 800
Db 2501 ACTCTGGCA---GAGACAGCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCTGGCAGC 2557
QY 801 ---ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db 2558 ACTGTCAACCCAGCTGGAGCTTATGCTTCAAGATCCCACTGCTCCATCCCGCAGAG 2617
QY 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db 2618 ATATGCAACAGCTAGTAGTCCCACTCAGCGGCAANTGACGTGGGATGTTAGCACAG 2677
QY 840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db 2678 MAGCTCTATGACCGGTACTGTAATTACTTTTGCACCAAGCGAGCCACCGGGTGTG 2737
QY 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAla 879
Db 2738 ATCTGGAGCTCTGGGAAGCTCTGACGAGGACGTGGGAGCTCAACACCTTGGCGAGT 2797
QY 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 2798 GCCTTGGAGGAGATGGCAAGAGTGTGCTGTGGCTGTGGCCACCGAGCGGGACTGC 2857

RESULT 13
US-10-087-684-3
; Sequence 3, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: MacDougall, John R.
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David J.
; APPLICANT: Grosse, William M.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie, J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Rastelli, Luca

```

; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles E.
; APPLICANT: Gangolli, Esha A.
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-214 CIP
; CURRENT APPLICATION NUMBER: US/10/087,684
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 3
; LENGTH: 2860
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (59)...(2857)
US-10-087-684-3

Alignment Scores:
Pred. No.: 1,72e-251 Length: 2860
Score: 2484.00 Matches: 493
Percent Similarity: 68.30% Conservative: 149
Best Local Similarity: 52.45% Mismatches: 250
Query Match: 51.89% Indels: 48
DB: 13 Gaps: 18

US-09-970-944-2 (1-899) x US-10-087-684-3 (1-2860)

QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyLeValLeuAlaLaTrp--- 19
Db 59 ATGGGGCCCGAGCGGAGCTCGGGCGCGCTGCTGCTGCGACTGCTGCTGCTGCGGAC 118
QY 20 -----LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAsnProValPro 36
Db 119 CCGAGGCTGAGCAAGCAGGCACTGATTCTGGCAGCGAGGTGCTCCCTGACTCTTCCG 178
QY 37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTrpIleVal 56
Db 179 TCAGCGCCAGCAGAGCGGCTGCTTCTCTGCGAGGAGCCACAGGAGCGCTTACATTGTG 238
QY 57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
Db 239 AAGAACAGCCCTGTGGAGCTTGGCTGCGCGCGCTTCCCGCCACACAGATCTACTTCAAG 298
QY 77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
Db 299 TGCACGCGGAGTGGGTGAGCCAGAACGACGCTCACAGAAAGGCGCTGATGAGGCC 358
QY 97 SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnValGluLysVal 116
Db 359 ACCGCGCTGCGGCTGCGGAGTGCAGATCGAGGTGTCGCGCAGCAGGTGGAGGAGCTC 418
QY 117 PheGlyLeuGluGluTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLys 136
Db 419 TTGCGCTGGAGGATTACTGTGGCCAGTGCCTGGGCTGAGGCTCGCGAGCACCACCAAG 478
QY 137 SerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeu 156

```

```

Db 479 AGTGGCCGAGCCTAGCTCGCATCGCTACCTGCGCAAGAACTTCGATCAGAGCCTCG 538
QY 157 AlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProGluGlyIle 176
Db 539 GGCAAGGAGGTGCCCTCGACCATGAGGTTCCTCTGCACTGCCGCCGCGGAGGGG 598
QY 177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196
Db 599 CCTGTGGCCGAGTGGATGGCTCAAGATGAGGATGTCATGACCCACCCAGGACACC 658
QY 197 AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216
Db 659 AACTTCCTGCTCACCATCGACCAACCTCATCTCCGCGAGCCGCTGCGGACACT 718
QY 217 AlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArgSerAlaAla 236
Db 719 GCCAACTATACCTGCTGGCCCAAGAACATCGTGGCCAAACGCGGAGCACCCTGCCACC 778
QY 237 ValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAla 256
Db 779 GTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGCAGAGTGGTCACTGCTCTCAAC 838
QY 257 SerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276
Db 839 CGCTGTGGCCGAGGCTGGCAGAGCGCACCCGAGCTGCACCAACCCGCTCCACTCAAC 898
QY 277 GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuVal 296
Db 899 GGAGGGGCTTCTCGAGGGCCAG--GCATTCCAGAGACCCGCTGCGACCATCTGCG 955
QY 297 SerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThr 316
Db 956 CCAGTCGATGGCGGCTGGAGCGAGTGGAGCAAGTGGTCCAGCTGCGACACTGAGTGTGCC 1015
QY 317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGln 336
Db 1016 CACTGGCGTAGCGGAGTGTCATGGCGGCCCCACCCAGAACGAGGCGCTGATGTCAGC 1075
QY 337 GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHis---SerAlaSer 355
Db 1076 GGGAGCGTCTCGACTCTAAGAACTGCACAGATGGGTGTGTCATCAACTGGAGGCTCA 1135
QY 356 GlyProGluAspValAlaIleuTyrValGlyLeu---IleAlaValAlaValCysLeuVal 374
Db 1136 GGG-----GATGCGCGCGCTGTATCGGGGCTCGTGGTGGCCATCTTGTGTCGTGCA 1189
QY 375 LeuLeuLeuLeuLeuValTyrCysArgLysLysGluGlyLeuAspSerAsp 394
Db 1190 ATCCTCATGGCGGTGGGGTGGTGTGTACCGCCCAACTGCCGTGACTTCGACACAGAC 1249
QY 395 ValAlaAspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSer 413
Db 1250 ATCACTGACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1309
QY 414 LysAlaAspAsnProHisLeuLeu-----ThrIleGlnProAspLeuSerThrThrThr 431
Db 1310 AGGCCAGTAACCCGAGCTCTACACCCCTCTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1369
QY 432 ---ThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeu 450
Db 1370 GGCATCTACCGGAGCCGCTGTATGCTCGCCTGCGAGGACTCC---ACGACAAAATCCCCA 1426
QY 451 ThrAsnGlyHisLeuLeuSerProLeuGlyGlyArgHisThrLeuHisHisSerSer 470
Db 1427 ACCAACTCTCTCTGCTGAGCCCTTACCAGCCTTAAAGTCAAGGTCAAGTCTACAGTCCAGC 1486
QY 471 ProThr-----SerGluAlaGluGluPheValSerArgLeuSerThr 484
Db 1487 ACCAGGGCTCTGGGCCAGGCTGCGAGATGGGCTGACCTGCTGGGGTCTTGTGCGCT 1546
QY 485 GlnAsnTyr-----PheArgSer----- 490

```


1547 GCACATACCTAGCGATTTCGCCGGGACACCCACTTCTGCACCTGCGCAGCGCCAGC 1606
 QY 491 -----LeuProArgGlyThrSerAsnMetThrTyrGlyThr 502
 Db 1607 CTCGGTTCCAGCAGCTCTTGGCCCTGCGCGAGACCCAGGAGCAGCGCTCAGCGGCACC 1666
 QY 503 PheAsnPheLeuGlyClyArgLeuMetIleProAsnThrGlyIleSerLeuIlePro 522
 Db 1667 TTTGGCTGCTGGTGGGAGCTCAGCATCCCGGCACAGGGGTGCTGTGTGGTGGCC 1726
 QY 523 ProAspAlaIleProArgGlyIleTyrGluIleTyrLeuThrLeuHisValProGlu 542
 Db 1727 AATGAGCCATTCCTCCAGGGCAAGTCTACGAGATGTATCTACTCATCAACAGGCAGAA 1786
 QY 543 AspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSerProIleValSerCys 561
 Db 1787 AGTACC---CTGCGCTTTTCAGAAAGGACCCAGACAGATATTGAGCCCTCGGTGACCTGT 1843
 QY 562 GlyProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGlu 581
 Db 1844 GGACCCACAGGCTCTCTGTGTGCCCGCCGTATCTCACCATGCCCCACTGTGCCGAA 1903
 QY 582 ProSerProAspSerTrpSerLeuArgLeuLeuLysGlnSerCysGluGlySerTrpGlu 601
 Db 1904 GTACGTGCCGTGACTGGATCTTTACGCTCAAGACCCAGGCCACAGGGCCACTGGAG 1963
 QY 602 GlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeuGlu 621
 Db 1964 ---GAGGTGCTGACCTCGATGAGGAGACCTGACACACACCTGTACTGCCAGCTGGAG 2020
 QY 622 AlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAla 641
 Db 2021 CCAGGCGCTGTACATCTGCTGGACGACCTGGCACCTTACGTGTTCAGCGCGAGTCC 2080
 QY 642 LeuSerValAlaAlaAlaLysArgLeuLeuLeuLeuPheAlaProValAlaCysThr 661
 Db 2081 TATTCGCTCAGCAGTCAAGCGCTCCAGCTGGCGCTCTTCGCCCGCCCTCTGCACC 2140
 QY 662 SerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGlu 681
 Db 2141 TCCTGGAGTACAGCTCCGGTCTACTGCTGGAGCAGCGCTGTAGCAGCTGAAGGAG 2200
 QY 682 ValValGlnLeuGlyGlnLeuGlyGlyGlnLeuIleGlnLeuProArgValLeuHis 701
 Db 2201 GTGCTGAGCTGGAGCGGACTTGGGCGGATCTTGTGGAGGAGCGAAACCGCTAATG 2260
 QY 702 PheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrp 721
 Db 2261 TTCAGGACAGTTTACCACAACTCGCGCTCTCCCTCCATGACCTCCGCCCATGCCATTGG 2320
 QY 722 LysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThr 741
 Db 2321 AGGAGCAAGCTGTCTGGCCAAATACAGGAGATCCCTTCTATCACATTTGAGTGGCAGC 2380
 QY 742 GlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeu 761
 Db 2381 CAGAGGCCCTCCACTGCATTTTCACCTGGAGAGGACAGCTGGCTCCACAGAGCTC 2440
 QY 762 AlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPhe 781
 Db 2441 ACCTGCAAGTCTCGTGGCGCAAGTGAAGGGGAGGGCCAGATATTCAGCTGCATACC 2500
 QY 782 AsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuSerGluAlaGly--- 800
 Db 2501 ACTCTGGCA---GAGACACCTGTGGCTCCCTGGACACTCTCTGCTCTGCCCTGGCAGC 2557
 QY 801 ---ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
 Db 2558 ACTGTACCCAGCTGGGACCTTATGCTCTCAAGATCCCACTGTCTCCATCCCGCAGAAG 2617
 QY 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
 Db 2618 ATATGCAACAGCTTAGATGCCGCCCACTACGCGGGCAATGACTGGCGGATGTTAGCAGC 2677

QY 840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
 Db 2678 AAGCTCTCTATGACCGGTACTGAATTAATTCGCCCAAGCGAGCCCGGCTGTG 2737
 QY 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
 Db 2738 ATCTCGGACCTCTGGGAAGCTCTGCAGCAGGAGCATGGGACCTCAACAGCCTGGCGAGT 2797
 QY 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
 Db 2798 GCCTTGAGGAGATGGCAAGAGTAGATGCTGGTGGCTGTGGCCACCGACGCGGAGCTGC 2857

RESULT 14

US-10-218-779-1
 ; Sequence 1, Application US/10218779
 ; Publication No. US20040029222A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Edinger, Shlomit
 ; APPLICANT: MacDougall, John
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Stone, David
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Alsobrook II, John
 ; APPLICANT: Lepley, Denise
 ; APPLICANT: Rieger, Daniel
 ; APPLICANT: Burgess, Catherine
 ; APPLICANT: Casman, Stacie
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Boldog, Ferenc
 ; APPLICANT: Li, Li
 ; APPLICANT: Fadigar, Muralidhara
 ; APPLICANT: Mishra, Vishnu
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Shenoy, Suresh
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Tchernev, Velizar
 ; APPLICANT: Vernet, Corine
 ; APPLICANT: Zerhusen, Bryan
 ; APPLICANT: Malyankar, Uriel
 ; APPLICANT: Guo, Xiaojia
 ; APPLICANT: Miller, Charles
 ; APPLICANT: Gangolli, Baha
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-214
 ; CURRENT APPLICATION NUMBER: US/10/218,779
 ; CURRENT FILING DATE: 2002-08-14
 ; PRIOR APPLICATION NUMBER: 60/253,834
 ; PRIOR FILING DATE: 2000-11-29
 ; PRIOR APPLICATION NUMBER: 60/250,-926
 ; PRIOR FILING DATE: 2000-11-30
 ; PRIOR APPLICATION NUMBER: 60/264,180
 ; PRIOR FILING DATE: 2001-01-25
 ; PRIOR APPLICATION NUMBER: 60/313,656
 ; PRIOR FILING DATE: 2001-08-20
 ; PRIOR APPLICATION NUMBER: 60/327,456
 ; PRIOR FILING DATE: 2001-10-05
 ; NUMBER OF SEQ ID NOS: 216
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 2860
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-218-779-1

Alignment Scores:
 Pred. No.: 1-72e-251 Length: 2860
 Score: 2484.00 Matches: 493
 Percent Similarity: 68.30% Conservative: 149
 Best Local Similarity: 52.45% Mismatches: 250
 Query Match: 51.89% Indels: 48

DB:	13	Gaps:	18
US-09-970-944-2 (1-899) x US-10-218-779-1 (1-2860)			
QY	1	MetAlaValArgProGlyLeuTyrProAlaLeuLeuGlyIleValLeuAlaIaIaTyrP----	19
DB	59	ATGGGGCGCCGAGCGAGGCTCGGGCGCGCTGCTGCTGCACACTGCTGCTGCTGGAC	118
QY	20	-----LeuArgGlySerGlyAlaGlnGln-----SerAlaThrValAlaAlaSerProValPro	36
DB	119	CCGAGGCTGAGCCACGACGACTGATCTCGGACGAGGTGCTCCCTCACTCTCCCG	178
QY	37	GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrIleVal	56
DB	179	TCAGCGCCAGCAGAGCGCTGCTTCTTCCTGCAGAGCCACAGACGCCCTACATTGTG	238
QY	57	LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys	76
DB	239	AAGAACAGACCTGTGAGGCTCGGCTCGCGCGCTTCCCGCCACACAGATCTACTTCAAG	298
QY	77	CysAsnGlyGluTyrValArgGlnValAspHisValIleGluArgSerThrAspGlySer	96
DB	299	TGCACCGCGAGTGGTTCAGCAGAACGACCGTACACAGAGAGCGCTGGATGAGGCC	358
QY	97	SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysVal	116
DB	359	ACCGTCTCGGGTGGCGAGGTGCAGATCGAGGTCTCGCGCAGCAGGTGGAGGACTC	418
QY	117	PheGlyLeuGluGluTyrTyrCysGlnCysValAlaTyrPheSerSerGlyThrThrLys	136
DB	419	TTTGGCTCGAGGATTACTTGGTGCAGTGGCTGCTGAGGTCCGCGGCACACCAAG	478
QY	137	SerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnProLeu	156
DB	479	AGTGGCGGAGCCTACCTCGCATCGCTACCTCGCGCAAGATTCGATCAGAGGCTCTG	538
QY	157	AlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProGluGlyIle	176
DB	539	GGCAAGGAGTGGCCCTGGACCATGAGTTCTCTGCACTGCGCGCCGCGGAGGGGTG	598
QY	177	ProProAlaGluValGluTyrLeuArgAsnGluAspLeuValAspProSerLeuAspPro	196
DB	599	CCTGTGGCGAGGTGAATGGTCAAGATGAGGATGTCATCGACCCACCCAGGACACC	658
QY	197	AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr	216
DB	659	AACTTCCTGCTCACCATCGACCAACCTCATCCGCGAGCGCGCTGCGGACACT	718
QY	217	AlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAla	236
DB	719	GCCAACTATACCTGGTGGCCAAAGAACATCGTGGCCAAACGCCGGAGCACCTGCCACC	778
QY	237	ValIleValTyrValAsnGlyTyrPheSerThrTyrThrGluTyrPheSerValCysSerAla	256
DB	779	GTCACTCGTCTAGCTGAATGGCGGCTGGTCCAGTGGCGAGAGTGGTCACTGCTCCAAAC	838
QY	257	SerCysGlyArgGlyTyrGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn	276
DB	839	CGCTGTGGCGGAGGCTGGCAGAGCGCACCCGGAGCTGCACCAACCCCGCTCCACTCAAC	898
QY	277	GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuVal	296
DB	899	GGAGGGGCTTCTGGAGGGCCAG---GCATTCCAGAGACCGCTGACACCACTCTGC	955
QY	297	SerValAspGlySerTyrPheSerProTyrPheSerAlaCysGlyLeuAspCysThr	316
DB	956	CCAGTCGATGGGCGTGGACGAGTGGAGCAAGTGGTCACTGCGAGCACTGAGTGTGCC	1015
QY	317	HisTyrPheArgGluCysSerAspProAlaProArgAsnGlyGluGluCysGln	336
DB	1016	CAGTGGCTAGCGGAGTGCATGGCGGCCCCACCCAGAGCGGCGCTGACTGCAGC	1075
QY	337	GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHis---SerAlaSer	355
DB	1076	GGGACGCTGCTGACTCTAAGAACTCCACAGATGGCTGTGCACTGGAGGGCTCA	1135
QY	356	GlyProGluAspValAlaLeuTyrValGlyLeu---IleAlaValAlaValCysLeuVal	374
DB	1136	CGG-----GATGGCGCGCTGATGCGGGGCTCGTGGTGCCCATCTTCGTGGTGGCA	1189
QY	375	LeuLeuLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAsp	394
DB	1190	ATCCCTCATCGCGTGGGGTGGTGTATCCGCCCACTGCGCGTGACTTCGACACAGAC	1249
QY	395	ValAlaAspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSer	413
DB	1250	ATCACTGACTCATCTGCTCCCTGACTGTGTGTTCACCCCGTCACTTTAAGACGGCA	1309
QY	414	LysAlaAspAsnProHisLeuLeu-----ThrIleGlnProAspLeuSerThrThrThr	431
DB	1310	AGGCCAGTAACCCGCGAGCTCTACACCCCTCTGTGCTCTCTGACCTGACAGCAGCGCC	1369
QY	432	---ThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeu	450
DB	1370	GGCATCTACCGCGGACCCGTGTATGCCCTGCGAGGACTCC---ACCGACAAATCCCCATG	1426
QY	451	ThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArgHisThrLeuHisHisSerSer	470
DB	1427	ACCAACTCTCTCTGCTGAGACCCCTTACCAGCCCTTAAGGTCAAGTCTACAGTCCAGC	1486
QY	471	ProThr-----SerGluAlaGluGluPheValSerArgLeuSerThr	484
DB	1487	ACCAAGGCTCTGGCGCAGGCTGGCAGATGGGCTGACCTGCTGGGGTCTTTGGCGCT	1546
QY	485	GlnAsnTyr-----PheArgSer-----	490
DB	1547	GGCACAATACCTAGCGATTTCGCCGGGACACCCACTTCTCGACCTGCGCAGCGCCAGC	1606
QY	491	-----LeuProArgGlyThrSerAsnMetThrTyrGlyThr	502
DB	1607	CTCGGTTCCAGCAGCTCTTGGGCTGCCCCGAGACCCAGGAGGAGCGTACAGCGCAC	1666
QY	503	PheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuLeuPro	522
DB	1667	TTTGGCTGCTGGTGGGAGGCTCAGCATCCCGGCACACAGGGTCTCAGCTTGTGGTCCC	1726
QY	523	ProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGlu	542
DB	1727	AATGAGCGCATTCGCCAGGGCAAGTCTACGAGATGTATCTACTCATCAACAAGCAGAA	1786
QY	543	AspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSerProIleValSerCys	561
DB	1787	AGTACC---CTGCCCTTTCAGAGGGACCCAGACAGTATTGAGCCCTCGGTGACCTGT	1843
QY	562	GlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGlu	581
DB	1844	GGACCCACAGGCTCTGCTGTCGCGCGCTCATCTCCTCACCATCCCGCTGTGCCGAA	1903
QY	582	ProSerProAspSerTyrSerLeuArgLeuLysLysGlnSerCysGluGlySerTyrGlu	601
DB	1904	GTCACTGCGCGTGAATGATCTTTTCAGCTCAAGACCCAGGCGCCACGAGGCGCTGGAG	1963
QY	602	GlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeuGlu	621
DB	1964	---GAGGTGGTGACCTCGATGAGGAGACCCCTGAACACACACCTGCTACTGCCAGCTGGAG	2020
QY	622	AlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAla	641
DB	2021	CCAGGCGCTGTACATCTCTGTGGACCACTGGGCACTACGTGTTACAGGGCGAGTCC	2080
QY	642	LeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAlaCysThr	661
DB	2081	TATTCCCTCAGCAGTCAAGCGGCTCAGTGGCGCTTTCGCCCGCGCTCTGTCACC	2140
QY	662	SerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGlu	681

```
Db 2141 TCCTGGAGTAGAGCTCGGGTCTACTGCTGGAGGACACGCCTGTAGACTGAGGAG 2200
Qy ValValGlnLeuGluysGlnLeuGlyGlnLeuLeuLeuLeuLeuLeuLeuHis 701
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2201 GTGCTGGAGTGGAGCGGACTCTGGCGGATACTTGGTGGAGGAGCGAAACCGCTAATG 2260
Qy PheLysAspSerTyHisAsnLeuArgLeuSerLleHisAspValProSerSerLeuTrp 721
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2261 TTCAAGGAGAGTTACCACAACTCGCCCTTCCTCCATGACCTCCCGATGCCCATTTGG 2320
Qy LysSerLysLeuLeuValSerTyGlnGlnLleProPheTyHisLleTrpAsnGlyThr 741
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2321 AGGAGCAAGCTGCTGCCAAATACCAGAGATCCCTTCTATACATTTGGATGGCAGC 2380
Qy GlnArgTyLysHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeu 761
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2381 CAGAAGGCCCTCACCTGCACCTTACCTCGAGAGGACAGACTTGGCCTCCACAGAGCTC 2440
Qy AlaCysLeuLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerLleAsnPhe 781
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2441 ACTGCAAGATCTGCTGGCGCAAGTGAAGGGAGGGGCGAGATATTCAGCTGCATACC 2500
Qy AsnLleThrLysAspThrArgPheAlaGluLeuAlaLeuLeuSerGluAlaGly--- 800
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2501 ACTCTGGCA--GAGACACCTGCTGGCTCCTGGACACACTCTGCTCTGCCCTGGCAGC 2557
Qy ---ValProAlaLeuValGlyProSerAlaPheLysLleProPheLeuLleArgGlnLys 819
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2558 ACTGTCAACACCAGCTGGGACCTTATGCTTCAAGATCCCACTGTCCATCCGCCAGAAG 2617
Qy IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2618 ATATGCAAGCCTAGATGCCCAACTACGCGGGCAATGACTGGCGGATTTAGCACAG 2677
Qy LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2678 AGCTTCTATGACCGGTACCTGAATTACTTTGCCAAAGCGAGCCCAACGGGTGTG 2737
Qy IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2738 ATCTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGACCTCAACAGCCTGGCGAGT 2797
Qy AlaValAlaGlyLeuGlnProAspAlaGlyLysPheThrValSerGluAlaGluCys 899
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2798 GCCTTGGAGAGATGGGCAAGAGTAGATGCTGCTGGCTGGCCACCGCGGGAGTGC 2857

RESULT 15
US-10-218-779-3
; Sequence 3, Application US/10218779
; Publication No. US2004002922A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit
; APPLICANT: MacDougall, John
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William
; APPLICANT: Alsobrook II, John
; APPLICANT: Lepley, Denise
; APPLICANT: Rieger, Daniel
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Sytek, Kimberly
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Patturajan, Meera
; APPLICANT: Shenoy, Suresh
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar
; APPLICANT: Vernet, Corine
```

```
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Malyankar, Uriel
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles
; APPLICANT: Gangolli, Esha
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-214
; CURRENT APPLICATION NUMBER: US/10/218,779
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,-926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-218-779-3

Alignment Scores:
Pred. No.: 1,72e-251 Length: 2860
Score: 2484.00 Matches: 493
Percent Similarity: 68.30% Conservative: 149
Best Local Similarity: 52.45% Mismatches: 250
Query Match: 51.89% Indels: 48
DB: 13 Gaps: 18

US-09-970-944-2 (1-899) x US-10-218-779-3 (1-2860)
Qy 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyLeValLeuAlaAlaTrp--- 19
Db 59 ATGGGGCGCGGAGCGAGCTCGGGCGGCTCTCTGCTGGCAGCTGCTCTCTCTGGGAC 118
Qy 20 -----LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAsnProValPro 36
Db 119 CCGAGGCTGAGCAACAGCAGGCACTGATTCTGGCAGCGAGTGTCTCCTGACTCTTCCCG 178
Qy 37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrlleVal 56
Db 179 TCAGCGCCAGCAGAGCGCTGCTCTCTCTGAGGAGCCAGGAGCGCTACATTGTG 238
Qy 57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnllePhePheLys 76
Db 239 AAGAACAAGCTGTGGAGCTTCGTCGCGGCGCTTCCCGCCACACAGATCTACTTCAAG 238
Qy 77 CysAsnGlyGluTrpValArgGlnValAspHisVallleGluArgSerThrAspGlySer 96
Db 299 TGCAACGGCGAGTGGGTGACGACCAAGCAGCAGTCCACAGGAGAGCGCTGGATGAGGCC 358
Qy 97 SerGlyGluProThrMetGluValArglleAsnValSerArgGlnGlnValGluLysVal 116
Db 359 ACCGGCTCGGGGTGGCGAGTGCAGATCGAGGTGTCGCGGAGCAGGTGGAGGAGCTC 418
Qy 117 PheGlyLeuGluGluTyrlTrpCysGlnCysValAlaTrpSerSerGlyThrThrLys 136
Db 419 TTTGGCTGGAGATTACTTGGTGCAGTGGCTGGCTCGGAGCTCCGAGCAGCACCACCAAG 478
Qy 137 SerGlnLysAlaTyrlleArglleAlaArgLeuArgLysAsnPheGluGlnGluProLeu 156
Db 479 AGTCGCGAGCGCTACGCTCGCATCGCTACCTCGCAAGAACTTCGATCAGGAGCCCTCTG 538
Qy 157 AlaLysGluValSerLeuGlnGlylleValLeuProCysArgProGluGlylle 176
Db 539 GGCAAGGAGGTGCCCTGGACCATGAGGTTCCTCTGAGTGCCTCGCGCGGAGGGGTG 598
```

177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196
 Db CTTGGCGGAGTGAATGGCTCAAGATGAGATGTCATGACCCCGAGGACACC 658
 197 AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216
 Db AACTTCCTGCTACCATCGACCAACAACCTCATCGCGCAGCGCGCTGTGGACACT 718
 217 AlaAsnTyrThrCysValAlaIleValAsnIleValAlaArgArgSerAlaSerAlaAla 236
 Db GCCAACTATACCTGCGTGGCCAGAACATCGTGGCCAAAGCCCGAGGACCATGCGCAC 778
 237 ValIleValTyrValAsnGlyGlyTyrSerThrTyrThrGluTyrSerValCysSerAla 256
 Db GTCATCGTCTACGTGAATGGCGCTGTCACAGCTGGGACAGTGGTCAACCTGCTCCAAC 838
 257 SerCysGlyArgGlyTyrGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276
 Db CGCTGTGGCGGAGGCTGGCAGAGCGCACCCCGACCTGCACCAACCCCGCTCCACTCAAC 898
 277 GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuVal 296
 Db GGAGGGGCGCTTTCGAGGGCCAG----GCATTCAGAGAGACCCGCTGCACCATCTGC 955
 297 SerValAspGlySerTyrProTyrSerTyrSerAlaCysGlyLeuAspCysThr 316
 Db CCAGTTCGATGGGCGTGGACGAGTGGAGCAAGTGTTCAGCCCTGCAGCACTGAGTGTGCC 1015
 317 HisTyrArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGln 336
 Db CACTGCGGTAGCGCGAGTGCATGGCGGCCGCCACCCAGACCGAGGCGGTGACTGCAGC 1075
 337 GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHis---SerAlaSer 355
 Db GGGACGCTGCTCACTCTAGAACTCCACAGATGGGCTGTGTGATCAACTGGAGGCGCTCA 1135
 356 GlyProGluAspValAlaLeuTyrValGlyLeu---IleAlaValAlaValCysLeuVal 374
 Db GGG-----GATGCGGCGGTGATGCGGGCTCGTGGTGGCCATCTTCGTGGTGGGCA 1189
 375 LeuLeuLeuValLeuIleLeuValTyrCysArgLysGluGlyLeuAspSerAsp 394
 Db ATTCCTCATGGCGGTGGGGTGGTGTGTACCGCGCAACTGCGCTGACTTCGACACAGAC 1249
 395 ValAlaAspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSer 413
 Db ATCACTGATCATCTGCTCCCTGACTGGTGGTTCACCCCGTCAACTTTAAGCGGCA 1309
 414 LysAlaAspAsnProHisLeuLeu-----ThrIleGlnProAspLeuSerThrThr 431
 Db AGGCGCCAGTAACCGGAGCTCTACACCCCTCTGTGCTCTCTGACTGACAGCCAGCGCC 1369
 432 ---ThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeu 450
 Db GGCATCTACCGCGGACCCGTGTATGCTGCTGCAGGACTCC---ACCGACAAAATCCCATG 1426
 451 ThrAsnGlyHisLeuLeuSerProLeuGlyGlyArgHisThrLeuHisSerSer 470
 Db ACCAACTCTCTGCTGAGCCCTTACCAGCCCTTAAGGTCAAGGTCTACAGCTCCAGC 1486
 471 ProThr-----SerGluAlaGluPheValSerArgLeuSerThr 484
 Db ACCACGGGCTCTGGGCCAGGCTGGCAGATGGGCTGACCTGCTGGGGGTCTTGGCGCT 1546
 485 GlnAsnTyr-----PheArgSer----- 490
 1547 GGCACATACCTAGCGATTTCGCCCGGACACCCCACTTCCTGCACCTGGCGAGCGCCAGC 1606
 491 -----LeuProArgGlyThrSerAsnMetThrTyrGlyThr 502
 1607 CTCGGTCCACAGAGCTCTTGGGCGCTGCCCCAGAGCCAGGAGGAGCGGTACGCGCAC 1666
 503 PheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePro 522

1667 TTTGGCTGCTGGTGGGAGCTCAGCATCCCGGCACAGGGGTGAGTTCGTGGTGCCC 1726
 523 ProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGlu 542
 Db AATGGAGCCATCCCGAGGCAAGTTCCTACGAGATGATCTACTCATCAAGAGGAGAA 1786
 543 AspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSerProIleValSerCys 561
 Db AGTACC---CTGCCGCTTTTCAGAAAGGAGCCACAGACAGTATTGAGCCCTCGTGACCTGT 1843
 562 GlyProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGlu 581
 Db GGAACCCAGGAGCTCTGTGTGCCCGCCGCTCATCTCACCATGCCCATCTGTGCCGAA 1903
 582 ProSerProAspSerTyrSerLeuArgLeuLysLysGlnSerCysGluGlySerTyrGlu 601
 Db GTCAGTGGCGGTGACTGGATCTTTCAGCTCAAGACCCAGGCCACCCAGGGCCACTGGAG 1963
 602 GlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeuGlu 621
 Db ---GAGGTGGTACCCCTGGATGAGGAGACCTTGAACACACCCCTGTCTACTGCCAGCTGGAG 2020
 622 AlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAla 641
 Db CCCAGGGCGCTGTCACATCTCTGTGACACGCTGGGCACCTAGTGTTCAGGGCGAGTCC 2080
 642 LeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThr 661
 Db TATTCGCTCAGCAGTCAAGCGGTCTCAGCTGGCGGTCTTCGCCCGCCCTCTGCACC 2140
 662 SerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGlu 681
 Db TCCCTGGAGTACAGCTCCGGTCTACTGCTGGAGGACACGCTGTACACACTGAAGGAG 2200
 682 ValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHis 701
 Db GTGCTGGAGCTGGAGGGAGCTCTGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATG 2260
 702 PheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTyr 721
 Db TTCAAGGACAGTTACCACAACTGCGCTCTCCCTCCATGACCTCCCATGCCATTGG 2320
 722 LysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTyrAsnGlyThr 741
 Db AGGACAAAGCTCTGCGCAATACCAGGAGATCCCTCTATCACAATTTGGAGTGGCAGC 2380
 742 GlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeu 761
 Db CAGAGGGCCCTCCATGACCTTTTCCCTGGAGAGGACAGCTTGGCTTCCACAGAGCTC 2440
 762 AlaCysLysLeuTyrValTyrGlnValGluGlyAspGlyGlnSerPheSerIleAsnPhe 781
 Db ACCTCAAGATCTGCTGCGGCAAGTGAAGGGAGGGCCAGATATTCAGCTGCATACC 2500
 782 AsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly--- 800
 Db ACTCTGGCA---GAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCTGGCAGC 2557
 801 ---ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
 Db ACTGTCAACCCAGCTGGGACCTTATGCTTCAAGATCCCACTGCTCCATCCGCGAGAG 2617
 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTyrArgThrLeuAlaGln 839
 Db ATATCAACAGCCTAGATGCCCCCAACTCAGGGGCAATGACTGGCGGATGTATGACAG 2677
 840 LysLeuHisLeuAspSerHisLeuSerPheAlaSerLysProSerProThrAlaMet 859
 Db AAGCTCTATGGACCGGTACCTGAATTAATTGTCACCAAGAGGAGGCGCCACGCGGTGTG 2737
 860 IleLeuAsnLeuTyrGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879

Db 2738 ATCTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGACCTCAACAGCCTGGCGAGT 2797
Qy 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
|||::: ::|||::: :::
Db 2798 GCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGC 2857

Search completed: October 6, 2004, 22:57:28
Job time : 1147 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 17:53:42 ; Search time 29 Seconds
(without alignments)
2981.937 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRGLWPAALLGIVLAWL.....AVAGLQPDAGLFTVSEAE 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	950.5	19.9	919	2 T32541	unc-5 protein - Ca
2	950.5	19.9	947	1 B4294	unc-5 protein, lon
3	280	5.8	1584	2 T00026	brain-specific ang
4	273.5	5.7	1172	1 TSHUP2	thrombospondin 2 p
5	273.5	5.7	1172	2 A42587	thrombospondin 2 p
6	270.5	5.7	1074	2 JCS928	semaphorin F precu
7	267.5	5.6	984	2 T00326	hypothetical prote
8	267.5	5.6	1522	2 T00028	brain-specific ang
9	263.5	5.5	1444	2 T18856	angiogenesis inhib
10	258	5.4	1572	2 T00027	brain-specific ang
11	248	5.2	1170	2 A40558	thrombospondin 1 p
12	246	5.1	1170	1 TSHUP1	thrombospondin 1 p
13	243.5	5.1	1178	1 A39804	thrombospondin pre
14	220.5	4.6	1651	2 T14160	transmembrane rece
15	217.5	4.5	1612	2 T30805	neural cell adhesi
16	211.5	4.4	1265	1 A37967	properdin - mouse
17	208.5	4.4	469	1 S29126	properdin precurs
18	206.5	4.3	254	2 T15952	hypothetical prote
19	202.5	4.2	437	2 S03478	hypothetical prote
20	198.5	4.1	788	2 T25061	properdin - mouse
21	188	3.9	423	2 T29549	hypothetical prote
22	185	3.9	1344	2 T14316	hypothetical prote
23	184	3.8	957	2 T15976	rig-1 protein - mo
24	183.5	3.8	1273	2 T42405	hypothetical prote
25	183.5	3.8	1863	2 S46217	sax-3 protein - Ca
26	176.5	3.7	860	2 T16892	protein-tyrosine-p
27	175.5	3.7	1736	2 A47747	hypothetical prote
28	170.5	3.6	1745	2 A46431	tight junction pro
29	169	3.5	934	1 A34372	tight junction-ass
					complement C6 prec

30 167.5 3.5 1907 2 S50893 protein-tyrosine-p
31 166 3.5 837 2 T00355 hypothetical prote
32 160.5 3.4 550 2 T47158 hypothetical prote
33 152.5 3.2 1499 2 I50212 protein-tyrosine-p
34 152 3.2 951 2 T00017 gene ADAMTS-1 prot
35 152 3.2 1501 2 I58148 protein-tyrosine-p
36 151 3.2 2165 2 T21371 hypothetical prote
37 150 3.1 1277 2 T30532 neural cell adhesi
38 149 3.1 654 2 T29247 hypothetical prote
39 148.5 3.1 946 1 A47299 ror-related recept
40 148 3.1 1091 2 A58532 glial cell membran
41 147 3.1 843 1 A27340 complement C7 prec
42 147 3.1 888 2 S23065 ufo protein - mous
43 146.5 3.1 584 1 C8HUA complement C8 alph
44 146 3.0 1051 2 A39712 kinase-like protei
45 145.5 3.0 2029 1 TDFFLK protein-tyrosine-p

ALIGNMENTS

RESULT 1

T32541

unc-5 protein - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 31-Jan-2000

C:Accession: T32541

R:Latreille, P.

submitted to the EMBL Data Library, December 1997

A:Description: The sequence of C. elegans cosmid B0273.

A:Reference number: Z21187

A:Accession: T32541

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-919 <LAT>

A:Cross-references: EMBL:AF036698; PIDN:AB88355.1; GSPDB:GNO0022; CESP:B0273.4a

A:Experimental source: strain Bristol N2; clone B0273

C:Genetics:

A:Gene: unc-5; CESP:B0273.4a

A:Map position: 4

A:Introns: 41/3; 108/1; 142/3; 201/1; 323/2; 553/3; 858/3

C:Superfamily: unc-5 protein; immunoglobulin homology; SH3 homology; thrombospondin type

Query Match 19.9%; Score 950.5; DB 2; Length 919;

Best Local Similarity 28.5%; Pred. No. 1.6e-61;

Matches 265; Conservative 165; Mismatches 375; Indels 125; Gaps 31;

QY 49 EPEDVIVVKNKPVLLVCKAVPATQIFFKNGEWVRQVDHVIER--STDGSSGEPTWEVRI 106

Db 9 QPKSGVIVNKKPLRLQCRANHATKIRYKCSKWID--DSRIEKLIGTDTSTGVGYIDASV 66

QY 107 NVYSROQVEKVFGLBYWCQCVAMSSSG-----TTKSQKAYIRIARLKNFEQELAKEYS 161

Db 67 DISRIDVDTSHTGDAFQCQCYA---SGDDQDQWASDVATVHLYMRKFLKSPVAQRVQ 123

QY 162 LEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVITREHSLVVRQARLADTANYTC 221

Db 124 EGTTLQLCPAESDPKAEALTWYKDGVVVQP--DANVIRASDGLSLSAARLSDSGNYTC 181

QY 222 VAKNTVARRRSASAIVVYVNGWSTWTEW-SVCSASCG-----RGWQKR 265

Db 182 EATNVANRKTDPVEQVYIVDGGWSENPWGTCTCHVDCLLRQHAHRIRDPHDVLPQHR 241

QY 266 SRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSDGSPWWSKWSACGLDCTHWSRECS 325

Db 242 TRTCNPAFLNDGEYCKGEEETRS-CKVPCKLDGWSWSWDSACSSCHYRTRACTV 300

QY 326 PAPRNGGECQGTDLDTNCTSDLCVHSASG--PDVALYVGLIAVAVCLVLLVLLV 383

Db 301 PPPMNGGQPCFGDDLMTOECPAQLCTADSSRIVSDTAVYGVSVASIFIVASFILAMF 360

QY 384 YCR-----KKEGLSDSVADS-----SILTSQGQP 407

Db 361 CCKRGNKSKPLKPKQNMSEKAGGIYSEPPGVRRLLEHGHQHTLLGKISCSQYFEP 420
 QY 408 VSIKPSKADNPHTLLTIQDPLSTTTTQGSILCPQDGPSPKFQLTNGHLLSPGLGGHRTLLH 467
 Db 421 PPL-----PHSTTLR---SGKSASFSGYSTFNAGSRAALIQECSSSSSGSGGKRTMLR 470
 QY 468 HSSPT-SEAEFVSRSLSTQNYFSLPRGTS-NMTYGTFFNLGLRLMIPNTGISLLIPDDA 525
 Db 471 TSSNCSDDNYATLYDMEDKSVLGLDTSQNTVAQAIDNSGARLSLSKSGARLIIVPELA 530
 QY 526 IPRGKIYEIYLTLLHKPEVDRLPLAGCQTLSPVSCGPGV-----LLTRPVILAMDH 578
 Db 531 VEGERM--LYLAVSDTLTQPHLKPIESALSPVIVIGQCDVMSAHDNLRPPVVSFRH 588
 QY 579 CGEPSP-DSWSLRKQKQCEGSEQVHLHGEAPSHLYYQLEASA-----CYVFT 629
 Db 589 CASTFPRDNWQFTL--YADEGSGWQKAVTIIGEENLNTNMFVQEPQKKGNDGFGWCHVMT 646
 QY 630 EQLGRLALYGEAL--SVAAKRLKLLLPAPVACTSLE--YNIRVYCLHDTHDALKKEVQL 695
 Db 647 YSLARLMLAGHPRNLSAARVHLAVFGPTMSAYRRPPELRVYCVPTGAAMESVWKQ 706
 QY 686 EKOLGGLIQEPR--VLHFKDSYHNLRLSIHDV-PSSLWKSLLVSYQEIPEFVHWNQTQ 742
 Db 707 ED--GSRLLCESNDFILNEKG---NLCICIEDVIPGSCDGPVEVVEISETQHRFV---AQ 758
 QY 743 RYLHCTFTLERSVPSSTDACKLWVQVEGDGQSGFSINFNITKDTFPAELALLASEAGVP 802
 Db 759 NGLHCSLKFRPKPEINGSQFSTRVIVYQKASSTPEVMW--EVSNEPELYDATSEEREKGSV 816
 QY 803 ALVGPAPKIPFLIRKQITSSLDPPCERGADWITLAKLHLDLHSHLFFFAKSP--SPTAMI 860
 Db 817 CV-----EPRLPFGVKDELARLLDMPNESHSDRWGLAKKHYRLOQFASFPDCSPTSL 872
 QY 861 LNLWEARHPFNGN-LSOLAAVAGLQOPDA 889
 Db 873 LDLWEASSSGSARVPDLTLQTLRVMGEPDA 902

RESULT 2
 unc-5 protein, long form - Caenorhabditis elegans
 N;Contains: unc-5 protein, short form
 C;Species: Caenorhabditis elegans
 C;Date: 30-Apr-1993 #sequence revision 28-Jul-1995 #text_change 05-Nov-1999
 C;Accession: B44294; T32540; A44294
 R;Leung-Hagetein, C.; Spence, A.M.; Stern, B.D.; Zhou, Y.; Su, M.W.; Hedgecock, E.M.;
 Cell 71, 289-299, 1992
 A;Title: UNC-5, a transmembrane protein with immunoglobulin and thrombospondin type 1 do
 A;Reference number: A44294; PMID:93046629; PMID:1384987
 A;Contents: variety Bergerac
 A;Accession: B44294
 A;Molecule type: DNA
 A;Residues: 1-947 <LEU>
 A;Cross-references: GB:S47168; NID:G258527; PID:AAB23867.1; PID:G258529
 A;Note: sequence extracted from NCBI backbone (NCBIN:116668, NCBIN:116670, NCBIN:116672,
 A;Note: authors translated the codon CTA for residue 642 as Val; sequence shown follows
 A;Note: mRNA lacking the first exon is equally prevalent
 R;Latrelle, P.
 submitted to the EMBL Data Library, December 1997
 A;Description: The sequence of C. elegans cosmid B0273.
 A;Reference number: 221187
 A;Accession: T32540
 A;Status: preliminary; translated from GB/EMBL/DBDJB
 A;Molecule type: DNA
 A;Residues: 1-947 <LAT>
 A;Cross-references: EMBL:AF036698; PIDN:AAB88356.1; GSPDB:GN00022; CESP:B0273.4b
 A;Experimental source: strain Bristol N2; clone B0273
 C;Genetics:
 A;Gene: unc-5
 A;Map position: 4
 A;Introns: 28/1; 69/3; 136/1; 170/3; 229/1; 351/2; 581/1; 886/3
 C;Function:

A;Description: required for guidance of pioneering axons and cells migrating dorsally al
 C;Superfamily: unc-5 protein; immunoglobulin homology; SH3 homology; thrombospondin type
 C;Keywords: alternative splicing; duplication; glycoprotein; receptor; transmembrane pro
 F;30-947/Product: unc-5 protein, short form #status predicted <ALT>
 F;46-116/Domain: immunoglobulin homology <IM1>
 F;153-211/Domain: immunoglobulin homology <IM2>
 F;229-300/Domain: thrombospondin type 1 repeat homology #status atypical <THR1>
 F;301-354/Domain: thrombospondin type 1 repeat homology <THR2>
 F;365-390/Domain: transmembrane #status predicted <TMW>
 F;512-559/Domain: SH3 homology <SH3>
 F;53-114,65-112,160-209/Disulfide bonds: #status predicted
 F;206/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.9%; Score 950.5; DB 1; Length 947;
 Best Local Similarity 28.5%; Pred. No. 1.7e-61;
 Matches 265; Conservative 165; Mismatches 375; Indels 125; Gaps 31;

QY 49 EPEDVIVKPKVLLVCKAVPATQIFKCNQWVRQVDHVIER--STDGSSSGEPTMEVRI 106
 Db 37 QPKSGVIRKPKLURLOCRANHATIKRYKSSKWID--DSRIEKLIGTSTSGVGYDASV 94
 QY 107 NVSRQQVEKVFGLIEYWCQCVAWSSSG-----TTKSQKAYIRIARLKNFQEPPLAKES 161
 Db 95 DISRIDVDTSGHVDAFQCQCYA--SGDDDDQDVASDVATVHLAYWKHFLKSPVQVRVQ 151
 QY 162 LEGIVLPCRPPEGPPIPAEVENLREDLVDPDLDPNVYITREHSLVVRQARLADTANYTC 221
 Db 152 EGTTLQPCAPESDPKAEITWYKGVVYQV--DANVIRASDGLSMAARSASDGSNYC 209
 QY 222 VAKNIVARRRSAAVIVVNGWSTWTEW-SVCSASCG-----RGWQKR 265
 Db 210 EATNVANSKTDVPEVQIIVDGGWSEWSPWIGTCHVDCPLLRQHAHRIRDPHDLPHQR 269
 QY 266 SRSCINPAPLNGAFCEGQNVHDRTVSSLLSVSDGWSWMSKWSACGLDCTHWRSECS 325
 Db 270 TRTCNNPAPLNDGEYCKGEBEEMTRS--CKVPCKLDGWSWSWMSACSSSCHRYRTRACT 328
 QY 326 PAPNCGECQCTDTRNCTSDLCVHSAG--PEDVALYVGLIAVAVCLVLLVLLV 383
 Db 329 PPPMNGGQPCFGDMLTQCPAQLCTADSSRIVISDTAVYGVSVASIFVASFILAIA 386
 QY 384 YCR-----KKEGLSDVADS-----SILTSQFOP 407
 Db 389 CCKRGNKSKPLKPKQNMSEKAGGIYSEPPGVRRLLEHGHQHTLLGKISCSQYFEP 448
 QY 408 VSIKPSKADNPHTLLTIQDPLSTTTTQGSILCPQDGPSPKFQLTNGHLLSPGLGGHRTLLH 467
 Db 449 PPL-----PHSTTLR---SGKSASFSGYSTFNAGSRAALIQECSSSSSGSGGKRTMLR 498
 QY 468 HSSPT-SEAEFVSRSLSTQNYFSLPRGTS-NMTYGTFFNLGLRLMIPNTGISLLIPDDA 525
 Db 499 TSSNCSDDNYATLYDMEDKSVLGLDTSQNTVAQAIDNSGARLSLSKSGARLIIVPELA 558
 QY 526 IPRGKIYEIYLTLLHKPEVDRLPLAGCQTLSPVSCGPGV-----LLTRPVILAMDH 578
 Db 559 VEGERM--LYLAVSDTLTQPHLKPIESALSPVIVIGQCDVMSAHDNLRPPVVSFRH 616
 QY 579 CGEPSP-DSWSLRKQKQCEGSEQVHLHGEAPSHLYYQLEASA-----CYVFT 629
 Db 617 CASTFPRDNWQFTL--YADEGSGWQKAVTIIGEENLNTNMFVQEPQKKGNDGFGWCHVMT 674
 QY 630 EQLGRLALYGEAL--SVAAKRLKLLLPAPVACTSLE--YNIRVYCLHDTHDALKKEVQL 685
 Db 675 YSLARLMLAGHPRNLSAARVHLAVFGPTMSAYRRPPELRVYCVPTGAAMESVWKQ 734
 QY 686 EKOLGGLIQEPR--VLHFKDSYHNLRLSIHDV-PSSLWKSLLVSYQEIPEFVHWNQTQ 742
 Db 735 ED--GSRLLCESNDFILNEKG---NLCICIEDVIPGSCDGPVEVVEISETQHRFV---AQ 786
 QY 743 RYLHCTFTLERSVPSSTDACKLWVQVEGDGQSGFSINFNITKDTFPAELALLASEAGVP 802
 Db 787 NGLHCSLKFRPKPEINGSQFSTRVIVYQKASSTPEVMW--EVSNEPELYDATSEEREKGSV 844

C:Superfamily: thrombospondin type 1 repeat homology
P:344-398/Domain: thrombospondin type 1 repeat homology <THR3>

	Query Match	5.6%	Score 267.5;	DB 2;	Length 984;
	Best Local Similarity	37.74;	Pred. No. 2.8e-111;		
	Matches 55;	Conservative 22;	Mismatches 54;	Indels 15;	Gaps 5;
Qy	220	TCVA-----KNIVARRRSAAAVIVVNGSGWSTWTEWSVCSASCGRGWQKRSRCTNPA	273		
	:	:	:	:	
	:	:	:	:	
	:	:	:	:	
Db	317	TCVSPYTHCGPLRESRVNCNTALCPVHGVEENSPWLSCSFTCGRGQRTTRSCT--P	374		
	:	:	:	:	
	:	:	:	:	
Qy	274	PLNGGAFCGQNVHDRVTSLLVSDGSFPKSNACGLDC---THWSREGSDAPRN	330		
	:	:	:	:	
	:	:	:	:	
Db	375	PQYGRPECEPETHHKPCNTALCPVDGQWDSWSQCSVTCNSNGTQQRSGCT--AAAH	432		
	:	:	:	:	
	:	:	:	:	
Qy	331	GGEECQGTDLDTNRCTSDLCVHSASG	356		
	:	:	:	:	
	:	:	:	:	
Db	433	GGEECRGFPAESRECYNPEC--TANG	456		
	:	:	:	:	
	:	:	:	:	

A:Protein resistance and abundance, range, growth rate, transmembrane, etc.
A:Experimental source: brain
C:Genetics:
A:Gene: GDB:BA13
A:Cross-references: GDB:9838090; OMIM:602684
A:Map position: 6q12-6q12
C:Superfamily: thrombospondin type 1 repeat homology
P:344-398/Domain: thrombospondin type 1 repeat homology <THR3>

Qy	331	GGECQGTDLTRNCTSDLCVHSAG	356
		: : : : :	
Db	433	GGSECRGFWAESRECYNPEC--TANG	456

QY 804 LVGPSAFK--IPFLIROKLISSLDPPCRGA--DWRTLAQKHLDSH----- 846
 Db 829 TVDGSFVSIVPFLV-PGIRYSVEVAASTGAGEGVKSEPFQLOSHGNGFVSPEDQVSLA 887
 QY 847 --LSFFASKPSPTA-----MTLNWEARHFPNGNLSOLAAAVAGLGOPDAGLFT 893
 Db 888 QQISDVVKQPAFIAGAACWILMVFSIWLVRHKKRN--GLSSTYAGIRKVPSTFT 944

RESULT 15
 T30805
 dutt1 protein - mouse
 N:Alternate names: transmembrane receptor protein Robol homolog
 C:Species: Mus musculus (house mouse)
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
 C:Accession: T30805
 R:Wu, M.C.; Lowe, N.; Fordham, R.; Rabbitts, P.
 submitted to the EMBL Data Library, July 1998
 A:Description: The mouse homologue of human DUTT1/H-robol gene: protein sequence and chr
 A:Reference number: Z20879
 A:Accession: T30805
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-1612 <WUM>
 A:Cross-references: EMBL:Y17793; NID:el329712; PID:el329713; PIDN:CAA76850.1
 A:Experimental source: brain
 C:Genetics:
 A:Gene: dutt1
 A:Map position: 16

Query Match 4.5%; Score 217.5; DB 2; Length 1612;
 Best Local Similarity 19.9%; Pred. No. 2.6e-07; Indels 355; Gaps 52;
 Matches 211; Conservative 133; Mismatches 359

QY 41 DLLHPLFVPEPDVYIVKPKVLLVCKAV--PATQIEFKCNGEWVROVDHVIERS----- 92
 Db 26 DFPFPRIVEHPSDLIVSKGEPATLCKAEGRPPTIEWYKGERV-ETDKDDPRSHMLLP 84
 QY 93 -----TDGSSGEPTEMEVRINVSROQVEKVFGLSEYWCQVAMSSSGTTKSKQAYIR 143
 Db 85 SGLSFLFLRIVHGRKSRPDEGVYI-----CVARNYLGEAVSHNASLE 125
 QY 144 IARLRNFFQEPFLAKVSLVLEQGLVLPDRPEGIPPAEVEWLRNEDLVPSLDENVYIT-R 202
 Db 126 VAILRDDFFQNSDVNVAVGEPAVMCEQPPRGHPTIISWKDGSPLD---DKDERITIR 182
 QY 203 EHSLVVRQARLADTANTYCVAKNIVARRRSASAAVIYYNGW-----STWTEWSV-- 253
 Db 183 GCKLMITYTRKSDAGKYVCVGTNNMGERESEVAELTVLERPSFVKRPSNLAVTVDDSAEF 242
 QY 254 -CSASCGRG-----WQK-----RSRSTNPAPLNGGAPCEGQNVH 287
 Db 243 KCEA---RGDPVTVWRKDDGELPKSRYEIRDHTLTKIRKVT--AGDMGSYTCVAENMV 297
 QY 288 DRTVSSLVSVDSG---NSPWSKWSACGLDCTHWSRECSDDAP-----RNGGEC-- 335
 Db 298 GKAEASATITVQEPHFVVKPRDQVVALGRVTI-FQCEATGNPQPAIFWREGSNLLPS 356
 QY 336 ----QGTDLDRNCTSDLCVHSASGPEVALYVGLIAVAVCLVLLVLLVLYCRKKEGL 391
 Db 357 YQPPQSSSRFSVSGTGLTITNVQ-RSDVGYI-----ICQTL----- 392
 QY 392 DSDVADSSILITSGFQFVSKPKADNPHELLTTPDLSTTTTYQGSJ---CPRQDGSPKPF 448
 Db 393 --NVA-GSIITKAY--LEVTDIADRPVPPVIRQGVNQTVANDGTLLILSCVATGSPAPTI 447
 QY 449 -----QLTNHGL---LSPLGG-GRHTLHSSPTSEAEFEVSRSLSTQNY 487
 Db 448 LWRKDGVLVSTQDSRIKQLESGLQIRYAKLGDGTGRYCTASTPSGEA----- 495
 QY 488 FRSLPRGTNMYGTNFTLGGRLMIPNTGISLLIP-----PDALPRKIVIEYTLHKPE- 542
 Db 496 -----TWSAY-----IEVQEFQVQVQPPRPPTDNLIFSAP-----SKPEV 530

QY 543 -DV---RLPLACQOTLLSPVSCGPPGVLLTRPVLAMDHCGEPPSPDSLSRLKKQSCGEG 598
 Db 531 TDVSKNTVTLSWQPNLNS-----GATPTSIIIEAFSHA---SGSSW-----QTAAB 573
 QY 599 SNEQDVHLHGEAPSHLYYCOLEASACVVFTE-----Q 631
 Db 574 NVKTFETFAIKGKPNAYILFLVRANAYGISDPQSISDPVKTDQVPPTSQGVDHKQVORE 633
 QY 632 LGRFALVGEALSVAATAAKRLKLLFAPVACTSLEYNIRVYCLHDTHTDALKEVVQLEKQLGG 691
 Db 634 LG-----NVVLHLHNPITILSSSSVEV-----HWTVDQOQSYIQ-----GY 668
 QY 692 QLIQEPRLVLFKDSYHNLRSLSDHVPSSLWKSKLL-----VSYQ--EIPFVHINWGTOR 743
 Db 669 KILYRPS-----GASHGESEWLVEVRTPTPKNSVWIPDLKRGVNYEIKARPPFNEFOGADS 724
 QY 744 YLHCTFTLERVSPSTSDLACKLWVQVEGDGQSFISNPN-ITKDTRFDEL-----LAL 795
 Db 725 EIKFAKTLEE-APSAPPRS--VTVSKNDGNGTALVTWQPPEDITIONGMVQYKVKWCLGN 781
 QY 796 ESEAGVPALVGPSPAFKIPFLIRQKIISLDDPPCR-----RGADWRTLAQKHLDS 845
 Db 782 ETKYHINKTVDGSTFSV-----VPSLVPGIRYSVEVAASGAGPGVKSEPFQTLQDS 834
 QY 846 H-----LSFFASKPSPTA-----MILNLWEARHFP--NGNLSOLA 878
 Db 835 HGNVPSPEDQVSLAQOISDVVRQPAFIAGIACWILMVFSIWLVRHKKRNLTLSTYA 894
 QY 879 A-----AVAGLGQPDAGLFTVSE 896
 Db 895 GIRKVPSTFTTPTVYQRGGEAVSSGGRP--GULNISE 930

Search completed: October 6, 2004, 18:20:08
 Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 15:50:42 ; Search time 19 Seconds

(without alignments)
2463.740 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRPLWFLGIVLAWL.....AVAGLGQPDAGLFTVSEAE 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	280	5.8	1584	1	BAL1_HUMAN
2	273.5	5.7	1172	1	TS22_HUMAN
3	273.5	5.7	1172	1	TS22_MOUSE
4	272.5	5.7	1077	1	SM5A_MOUSE
5	270.5	5.7	1074	1	SM5A_HUMAN
6	269	5.6	1170	1	TS22_BOVIN
7	267.5	5.6	1522	1	BA13_HUMAN
8	266.5	5.6	1173	1	TS21_XENLA
9	258.5	5.4	1093	1	SM5B_MOUSE
10	258	5.4	1572	1	BA12_HUMAN
11	252	5.3	1093	1	SM5B_HUMAN
12	248	5.2	1170	1	TS21_MOUSE
13	246	5.1	1170	1	TS21_HUMAN
14	243.5	5.1	1178	1	TS22_CHICK
15	243	5.1	1170	1	TS21_BOVIN
16	208.5	4.4	469	1	PRO1_HUMAN
17	205.5	4.3	470	1	PROP_CAVPO
18	202.5	4.2	437	1	PROP_MOUSE
19	186	3.9	1266	1	NGCA_CHICK
20	181.5	3.8	867	1	SSPO_BOVIN
21	178.5	3.7	1223	1	AT14_HUMAN
22	175.5	3.7	1736	1	ZOL1_HUMAN
23	172	3.6	630	1	AT34_RAT
24	170.5	3.6	1745	1	ZOL1_MOUSE
25	169	3.5	934	1	C06_HUMAN
26	167.5	3.5	905	1	AT58_MOUSE
27	166	3.5	837	1	AT54_HUMAN
28	162.5	3.4	587	1	C08B_ONCMY
29	160.5	3.4	562	1	AT15_MOUSE
30	160.5	3.4	967	1	AT51_MOUSE
31	160.5	3.4	1077	1	AT10_HUMAN
32	159.5	3.3	1224	1	AT16_HUMAN
33	156.5	3.3	860	1	AT56_HUMAN
34	156	3.3	950	1	AT15_HUMAN
35	155	3.2	997	1	AT57_HUMAN
36	154.5	3.2	1593	1	AT12_HUMAN
37	154	3.2	1213	1	AT52_MOUSE
38	153.5	3.2	1205	1	AT53_HUMAN
39	152	3.2	967	1	AT51_RAT
40	152	3.2	968	1	AT51_MOUSE
41	151.5	3.2	890	1	AT58_HUMAN
42	151.5	3.2	1095	1	AT17_HUMAN
43	151	3.2	930	1	AT55_HUMAN
44	150	3.1	1277	1	CAML_FUGRU
45	149	3.1	140	1	C08B_RAT

ALIGNMENTS

RESULT 1

ID	BAL1_HUMAN	STANDARD;	PRT;	1584 AA.
AC	O14514;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Brain-specific angiogenesis inhibitor 1 precursor.			
GN	BAL1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Fetal brain;			
RX	MEDLINE=98054121; PubMed=9393972;			
RA	Nishimori H., Shiratsuchi T., Urano T., Kimura Y., Kiyono K.,			
RA	Tatsumi K., Yoshida S., Ono M., Kuwano M., Nakamura Y., Tokino T.;			
RT	"A novel brain-specific p53-target gene, BAL1, containing			
RT	thrombospondin type 1 repeats inhibits experimental angiogenesis.";			
RL	Oncogene 15:2145-2150(1997).			
RN	[2]			
RP	INTERACTION WITH BAP1.			
RX	MEDLINE=98321173; PubMed=9647739;			
RA	Shiratsuchi T., Futamura M., Oda K., Nishimori H., Nakamura Y.,			
RA	Tokino T.;			
RT	"Cloning and characterization of BAL-associated protein 1: a PDZ			
RT	domain-containing protein that interacts with BAP1.";			
RL	Biochem. Biophys. Res. Commun. 247:597-604(1998).			
CC	!- FUNCTION: LIKELY TO BE A POTENT INHIBITOR OF ANGIOGENESIS IN			
CC	BRAIN AND MAY PLAY A SIGNIFICANT ROLE AS A MEDIATOR OF THE P53			
CC	SIGNAL IN SUPPRESSION OF GLOBLASTOMA. MAY FUNCTION IN CELL			
CC	ADHESION AND SIGNAL TRANSDUCTION IN THE BRAIN.			
CC	!- SUBUNIT: INTERACTS WITH BAP1.			
CC	!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. LIKELY TO BE			
CC	CONCENTRATED AT CELL-CELL ADHESION SITES.			
CC	!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN BRAIN. REDUCED OR NO			
CC	EXPRESSION IS OBSERVED IN SOME GLOBLASTOMA CELL LINES AND CANCER			
CC	TISSUES.			
CC	!- INDUCTION: By p53.			
CC	!- DOMAIN: THE TSP1 REPEATS INHIBIT IN VIVO ANGIOGENESIS IN RAT			
CC	CORNEA INDUCED BY BFGF.			
CC	!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.			
CC	!- SIMILARITY: Contains 5 TSP type-1 domains.			
CC	!- SIMILARITY: Contains 1 GPS domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AB005297; BAA3647.1; --			

O8te58 homo sapien
Q9ukp4 homo sapien
P58397 homo sapien
Q8c9w3 m adamts-2
O15072 homo sapien
Q9wuq1 rattus norv
P97857 mus musculu
Q9up79 homo sapien
Q8te56 homo sapien
Q9una0 homo sapien
Q98902 fugu rubrip
P55314 rattus norv

FT	DISULFID	397	430	BY SIMILARITY.
FT	DISULFID	408	415	BY SIMILARITY.
FT	DISULFID	449	486	BY SIMILARITY.
FT	DISULFID	453	491	BY SIMILARITY.
FT	DISULFID	464	476	BY SIMILARITY.
FT	DISULFID	506	543	BY SIMILARITY.
FT	DISULFID	510	548	BY SIMILARITY.
FT	DISULFID	521	533	BY SIMILARITY.
FT	DISULFID	553	584	BY SIMILARITY.
FT	DISULFID	558	574	BY SIMILARITY.
FT	DISULFID	577	588	BY SIMILARITY.
FT	DISULFID	594	610	BY SIMILARITY.
FT	DISULFID	601	619	BY SIMILARITY.
FT	DISULFID	622	646	BY SIMILARITY.
FT	DISULFID	652	665	BY SIMILARITY.
FT	DISULFID	659	678	BY SIMILARITY.
FT	DISULFID	680	691	BY SIMILARITY.
FT	DISULFID	707	715	BY SIMILARITY.
FT	DISULFID	720	740	BY SIMILARITY.
FT	DISULFID	756	776	BY SIMILARITY.
FT	DISULFID	779	799	BY SIMILARITY.
FT	DISULFID	815	835	BY SIMILARITY.
FT	DISULFID	838	858	BY SIMILARITY.
FT	DISULFID	876	896	BY SIMILARITY.
FT	DISULFID	912	932	BY SIMILARITY.
FT	DISULFID	948	1169	BY SIMILARITY.
FT	CARBOHYD	151	151	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	316	316	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	330	330	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	457	457	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	584	584	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	710	710	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1069	1069	N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQ	SEQUENCE	1172 AA;	129911 MW;	7CE8B4E8599822AB CRC64;

Query Match 5.7%; Score 273.5; DB 1; Length 1172;
Best Local Similarity 36.1%; Pred. No. 7.9e-13;
Matches 57; Conservative 22; Mismatches 70; Indels 9; Gaps 4;

QY	209	ROARLADTANYTCVAKNIVARRRS-ASAAVIVVNGGWSWTETWSVCSASCGRGWQKRSR	267
DB	403	QGRSCDVTNLCGLPSIQTRICSLGKCDTRIQNGGWSHSPSSCSVTCGVGNVTRIR	462
QY	268	SCTNPAPLNGAPFCEQNVHVRTVSSLLVSVDGWSPWKWSACGLDCT---HWRSECS	324
DB	463	LCNSPPVQPGGKNCKSGRETQPCQDPIDGRWSPWSPSACTVTCAGGIRRSRVCN	522
QY	325	DPAPRNGECCG--TD---LDRNCTGSDLCVHSASGP	357
DB	523	SPEPYGGKDKDGVDTHEHMCNKRSCPIDGCLSNPCFP	560

RESULT 4
SM5A MOUSE
ID SM5A_MOUSE STANDARD; PRT; 1077 AA.

AC	Q62217;
DT	30-MAY-2000 (Rel. 39, Created)
DT	30-MAY-2000 (Rel. 39, Last sequence update)
DT	28-FEB-2003 (Rel. 41, Last annotation update)
DE	Semaphorin 5A precursor (Semaphorin F) (Sema F).
GN	SEMA5A OR SEMAF OR SEMF.
OS	Mus musculus (Mouse)
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RC	SEQUENCE FROM N.A.
RP	STRAIN=NMRI.
RX	MEDLINE=96414430; PubMed=8817451;
RA	Adams R.H.; Bezzi H.; Puschel A.W.;
RT	"A novel class of murine semaphorins with homology to thrombospondin
RT	is differentially expressed during early embryogenesis.";
RL	Mech. Dev. 57:33-45(1996).


```

FT DOMAIN 784 839 TSP TYPE-1 5.
FT DOMAIN 841 896 TSP TYPE-1 6.
FT DOMAIN 897 944 TSP TYPE-1 7.
FT CARBOHYD 142 144 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 227 227 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 323 323 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 367 367 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 591 591 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 717 717 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 933 933 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 56 56 A -> V (IN REF. 2).
FT CONFLICT 149 149 A -> T (IN REF. 2).
FT CONFLICT 382 382 V -> M (IN REF. 2).
FT CONFLICT 494 494 S -> R (IN REF. 2).
SQ SEQUENCE 1074 AA; 120570 MW; E53DB763CBE29407 CRC64;

Query Match 5.7%; Score 270.5; DB 1; Length 1074;
Best Local Similarity 42.4%; Pred. No. 1.2e-12;
Matches 50; Conservative 13; Mismatches 52; Indels 3; Gaps 1;

QY 241 VNGGWTWTSVCSACRGWKRSCNPNAPANGGAFCEGQNVHDTVSSLLYSVDG 300
DB 783 VNGAWSAWTSWSQCRSDCRGIRNRKVCNPNPKYGGMPCLGFSLEYQRCNTLPCFVDG 842

QY 301 SWSPEKWSAGCLDGC--THWRSGCDPAPRNGGEECCQGLDTRNCTSDLCVHGS 355
DB 843 VWSGSPFWTKSATCGGHHYMRTRSCSNPAPAGGIDICLGLHTEBALCNTQPCFWS 900

RESULT 6
TSP2_BOVIN STANDARD; PRT; 1170 AA.
AC Q95116; Q28180;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thrombospondin 2 precursor (Corticotropin-induced secreted protein)
DE (CIS)
GN THS2 OR TSP2 OR TSP-2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Danik M., Chinn A., Lafeuillade M., Keramidas M., Aguesse-Germon S.,
RA Penhoat A., Chen H., Mosher D., Chambaz E.M., Feige J.J.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-522 FROM N.A.
RX MEDLINE=9631130; PubMed=8698834;
RA Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,
RA Feige J.J.;
RT "Opposite regulation of thrombospondin-1 and corticotropin-induced
RT secreted protein/thrombospondin-2 expression by adrenocorticotrophic
RT hormone in adrenocortical cells.";
RL J. Cell. Physiol. 167:164-172(1996).
RN [3]
RP SEQUENCE OF 318-831 FROM N.A.
RC TISSUE=Aortic endothelium;
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;
RT "Cloning and sequencing of bovine thrombospondin stimulatory effect of
RT TGF-beta.";
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin and type V collagen.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.

```

```

CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 1 WFEC domain.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 7 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X96540; CAA65385.1; -
CC DB EMBL; X87620; CAA60952.1; -
CC DR HSSP; P00740; 1EDM.
CC DR InterPro; IPR001881; EGF_Ca.
CC DR InterPro; IPR006209; EGF_like.
CC DR InterPro; IPR006210; IEGF.
CC DR InterPro; IPR000884; TSP1.
CC DR InterPro; IPR008085; TSP 1.
CC DR InterPro; IPR003367; tsp 3.
CC DR InterPro; IPR008859; TSPC.
CC DR InterPro; IPR003129; TSPN.
CC DR InterPro; IPR001007; VWF_C.
CC DR Pfam; PF00008; EGF; 1.
CC DR Pfam; PF00093; vwc; 1.
CC DR PRINTS; PR01705; TSP1REPEAT.
CC DR SMART; SM00181; EGF; 3.
CC DR SMART; SM00209; TSP1; 3.
CC DR SMART; SM00210; TSPN; 1.
CC DR SMART; SM00214; VWC; 1.
CC DR PROSITE; PS00022; EGF_1; FALSE_NEG.
CC DR PROSITE; PS01186; EGF_2; 1.
CC DR PROSITE; PS50026; EGF_3; 2.
CC DR PROSITE; PS50092; TSP1; 3.
CC DR PROSITE; PS01208; VWF_1; 1.
CC DR PROSITE; PS50184; VWF_2; 1.
CC KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW EGF-like domain; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170
FT DOMAIN 19 215 THROMBOSPONDIN 2.
FT DOMAIN 19 232 TSP N-TERMINAL.
FT DOMAIN 318 375 HEPARIN-BINDING (POTENTIAL).
FT DOMAIN 379 429 WFEC.
FT DOMAIN 435 490 TSP TYPE-1 1.
FT DOMAIN 492 547 TSP TYPE-1 2.
FT DOMAIN 547 587 TSP TYPE-1 3.
FT DOMAIN 588 645 EGF-LIKE 1.
FT DOMAIN 646 690 EGF-LIKE 2.
FT DOMAIN 723 758 EGF-LIKE 3.
FT DOMAIN 759 781 TSP TYPE-3 1.
FT DOMAIN 782 817 TSP TYPE-3 2.
FT DOMAIN 818 840 TSP TYPE-3 3.
FT DOMAIN 841 878 TSP TYPE-3 4.
FT DOMAIN 879 914 TSP TYPE-3 5.
FT DOMAIN 915 950 TSP TYPE-3 6.
FT DOMAIN 951 1170 TSP TYPE-3 7.
FT SITE CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 266 266 INTERCHAIN (PROBABLE).
FT DISULFID 270 270 INTERCHAIN (PROBABLE).
FT DISULFID 391 423 BY SIMILARITY.
FT DISULFID 395 428 BY SIMILARITY.
FT DISULFID 406 413 BY SIMILARITY.
FT DISULFID 447 484 BY SIMILARITY.
FT DISULFID 451 489 BY SIMILARITY.

```

```

FT DISULFID 462 474 BY SIMILARITY.
FT DISULFID 504 541 BY SIMILARITY.
FT DISULFID 508 546 BY SIMILARITY.
FT DISULFID 519 531 BY SIMILARITY.
FT DISULFID 551 562 BY SIMILARITY.
FT DISULFID 556 572 BY SIMILARITY.
FT DISULFID 575 586 BY SIMILARITY.
FT DISULFID 592 608 BY SIMILARITY.
FT DISULFID 599 617 BY SIMILARITY.
FT DISULFID 620 644 BY SIMILARITY.
FT DISULFID 650 663 BY SIMILARITY.
FT DISULFID 657 676 BY SIMILARITY.
FT DISULFID 678 689 BY SIMILARITY.
FT DISULFID 705 713 BY SIMILARITY.
FT DISULFID 718 738 BY SIMILARITY.
FT DISULFID 754 774 BY SIMILARITY.
FT DISULFID 777 797 BY SIMILARITY.
FT DISULFID 813 833 BY SIMILARITY.
FT DISULFID 836 856 BY SIMILARITY.
FT DISULFID 874 894 BY SIMILARITY.
FT DISULFID 910 930 BY SIMILARITY.
FT DISULFID 946 1167 BY SIMILARITY.
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 316 316 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 582 582 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 708 708 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 936 936 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1067 1067 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 535 535 A -> V (IN REF. 3).
FT CONFLICT 748 748 S -> T (IN REF. 3).
SQ SEQUENCE 1170 AA; 129862 MW; 9CF1FBP5B89A051 CRC64;

Query Match 5.6%; Score 269; DB 1; Length 1170;
Best Local Similarity 35.6%; Pred. No. 1.8e-12;
Matches 52; Conservative 21; Mismatches 69; Indels 4; Gaps 2;

QY 209 ROARLADTANTVCVAKNIVARRRSASAA-VIVVNGWSTWTEWSVCASCGRGWQKRSR 267
DB 401 QRGSCDVTSNTCLGFSIQTRACSLGRCDHRIQDGGWHSWPSWSSCVTCGVGNVTRIR 460
QY 268 SCTNPAPLNGGAFCEGQNVHRTVSSLLVSVDSWSPWSKWSACGLDCT--HWRSRECS 324
DB 461 LCNSPVPQMGGRSCCKSGRETKACQPGPCPDVGRSPWSPWACVITCAGGIRRETRVCN 520
QY 325 DPAPNGGEBECQGLDTRNCTSDLC 350
DB 521 SPEPQHGKDCVGGAKEQCMNRKSC 546

RESULT 7
ID3_HUMAN
AC O60242; O60297; STANDARD; PRT; 1522 AA.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Brain-specific angiogenesis inhibitor 3 precursor.
GN BAI3 OR KIAA0550.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Fetal brain;
RA MEDLINE=98194217; PubMed=9533023;
RT "Cloning and characterization of BAI2 and BAI3, novel genes homologous
RT to brain-specific angiogenesis inhibitor 1 (BAI1).";
RL Cytogenet. Cell Genet. 79:103-108(1997).
RN [2]

RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98290545; PubMed=9628581;
RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RL code for large proteins in vitro.";
RNA Res. 5:31-39(1998).
RN [3]
RP SEQUENCE FROM N.A., AND REVISIONS TO 643-665 AND C-TERMINUS.
RX MEDLINE=22158633; PubMed=12168954;
RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
RT "Construction of expression-ready cDNA clones for KIAA genes: manual
RT curation of 330 KIAA cDNA clones.";
RNA Res. 9:99-106(2002).
CC -!- FUNCTION: MIGHT BE INVOLVED IN ANGIOGENESIS INHIBITION AND
CC SUPPRESSION OF GLIOBLASTOMA.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN BRAIN. ALSO DETECTED IN
CC HEART. REDUCED EXPRESSION IS OBSERVED IN SOME GLIOBLASTOMA CELL
CC LINES.
CC -!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 4 TSP type-1 domains.
CC -!- SIMILARITY: Contains 1 GPS domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AB005299; BAA25363.1; -.
CC EMBL; AB011122; BAA25476.2; ALT_INIT.
CC PIR; T00028; T00028.
CC GENE; HGNC:945; BAI3.
CC MIM; 602684; -.
CC
CC InterPro; IPR000859; CUB.
CC InterPro; IPR000832; GPCR secretin.
CC InterPro; IPR001879; hormn_receptor.
CC InterPro; IPR000203; PKD_cys_rich.
CC InterPro; IPR000884; TSP1.
CC Pfam; PF00002; 7tm_2; 1.
CC Pfam; PF01825; GPS; 1.
CC Pfam; PF02793; HRM; 1.
CC Pfam; PF00090; tsp_1; 4.
CC SMART; SM00303; GPS; 1.
CC SMART; SM00008; Hormr; 1.
CC SMART; SM00209; TSP1; 4.
CC PROSITE; PS01180; CUB; 1.
CC PROSITE; PS02221; GPS; 1.
CC PROSITE; PS00649; G_PROTEIN_RECP_F2_1; FALSE_NEG.
CC PROSITE; PS00650; G_PROTEIN_RECP_F2_2; FALSE_NEG.
CC PROSITE; PS02227; G_PROTEIN_RECP_F2_3; 1.
CC PROSITE; PS02261; G_PROTEIN_RECP_F2_4; 1.
CC PROSITE; PS00092; TSP1; 4.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
CC Repeat.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 1522 BRAIN-SPECIFIC ANGIOGENESIS INHIBITOR 3.
FT DOMAIN 25 880 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 881 901 1 (POTENTIAL).
FT DOMAIN 902 910 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 911 931 2 (POTENTIAL).
FT DOMAIN 932 939 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 940 960 3 (POTENTIAL).
FT DOMAIN 961 981 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 982 1002 4 (POTENTIAL).
FT DOMAIN 1003 1023 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1024 1044 5 (POTENTIAL).

```

```

FT DOMAIN 1045 1098 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1099 1119 6 (POTENTIAL).
FT DOMAIN 1120 1125 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1126 1146 7 (POTENTIAL).
FT DOMAIN 1147 1522 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 30 159 CUB.
FT DOMAIN 291 343 TSP TYPE-1 1.
FT DOMAIN 345 398 TSP TYPE-1 2.
FT DOMAIN 400 453 TSP TYPE-1 3.
FT DOMAIN 455 508 TSP TYPE-1 4.
FT DOMAIN 816 868 GPS.
FT DOMAIN 942 945 POLY-THR.
FT DOMAIN 1173 1176 POLY-SER.
FT CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 337 337 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 540 540 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 779 779 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 828 828 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 937 937 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1522 AA; 171490 MW; D22D0A5D4BB62502 CRC64;

Query Match 5.6%; Score 267.5; DB 1; Length 1522;
Best Local Similarity 37.7%; Pred. No. 3.3e-12;
Matches 55; Conservative 22; Mismatches 54; Indels 15; Gaps 5;

QY 220 TCVA-----KNIVARRRSAAVIVYNGWSTWTSVCSASCGRWQKRSRCTNPA 273
Db 317 TCVSPYGTCHSGFLRESRCNNATLCPVHGVMWESWSPSLCSFTCGRGQRTSRCT--P 374
QY 274 PLNGAFCEQNVHDTVSLVSDGSPSKWSACGLDC---THWESRCSQDPAPRN 330
Db 375 PQYGGPRCPGEPETHHPCNIALCPVDGQWQWESWQSCSVTCNSGTFQKRSROCT--AAAH 432
QY 331 GBECEGOTDLDTNCTSDLCVHSASG 356
Db 433 GSECEGPAWESRECVNPEC--TANG 456

RESULT 8
TSP1_XENLA STANDARD; PRT; 1173 AA.
AC P35448;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thrombospondin 1 precursor.
GN THBS1 OR TSP1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;
RL Submitted (XXX-1993) to the EMBL/GenBank/DDAJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIB/beta-3 (By similarity).
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 1 WFCC domain.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.

```

```

CC CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC -----
CC DR EMBL; L04278; -; NOT ANNOTATED_CDS.
CC DR HSSP; P00740; 1EDM.
CC DR InterPro; IPR001881; EGF_Ca.
CC DR InterPro; IPR006209; EGF-like.
CC DR InterPro; IPR006210; IEGF.
CC DR InterPro; IPR000884; TSP1.
CC DR InterPro; IPR008085; TSP 1.
CC DR InterPro; IPR003367; tsp 3.
CC DR InterPro; IPR008859; TSPC.
CC DR InterPro; IPR003129; TSPN.
CC DR InterPro; IPR001007; VWF_C.
CC DR Pfam; PF00008; EGF; 2.
CC DR Pfam; PF00090; tsp 1; 3.
CC DR Pfam; PF02412; tsp 3; 13.
CC DR Pfam; PF05735; TSPC; 1.
CC DR Pfam; PF02210; TSPN; 1.
CC DR Pfam; PF00093; VWC; 1.
CC DR PRINTS; PR01705; TSP1REPEAT.
CC DR SMART; SM00181; EGF; 2.
CC DR SMART; SM00209; TSP1; 3.
CC DR SMART; SM00210; TSPN; 1.
CC DR SMART; SM00214; VWC; 1.
CC DR PROSITE; PS00022; EGF_1; FALSE_NEG.
CC DR PROSITE; PS01186; EGF 2; 1.
CC DR PROSITE; PS00026; EGF 3; 2.
CC DR PROSITE; PS00092; TSP1; 3.
CC DR PROSITE; PS01208; WFCC 1; 1.
CC DR PROSITE; PS0184; WFCC 2; 1.
CC KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
CC EGF-like domain; Signal.
FT SIGNAL 1 22
FT CHAIN 23 1173 THROMBOSPONDIN 1.
FT DOMAIN 23 235 HEPARIN-BINDING (POTENTIAL).
FT DOMAIN 23 224 TSP N-TERMINAL.
FT DOMAIN 319 376 WFCC.
FT DOMAIN 382 432 TSP TYPE-1 1.
FT DOMAIN 438 493 TSP TYPE-1 2.
FT DOMAIN 495 550 TSP TYPE-1 3.
FT DOMAIN 550 590 EGF-LIKE 1.
FT DOMAIN 591 648 EGF-LIKE 2.
FT DOMAIN 649 693 EGF-LIKE 3.
FT DOMAIN 726 761 TSP TYPE-3 1.
FT DOMAIN 762 784 TSP TYPE-3 2.
FT DOMAIN 785 820 TSP TYPE-3 3.
FT DOMAIN 821 843 TSP TYPE-3 4.
FT DOMAIN 844 881 TSP TYPE-3 5.
FT DOMAIN 882 917 TSP TYPE-3 6.
FT DOMAIN 918 953 TSP TYPE-3 7.
FT DOMAIN 954 1173 C-TERMINAL.
FT SITE 929 931 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 394 426 BY SIMILARITY.
FT DISULFID 398 431 BY SIMILARITY.
FT DISULFID 409 416 BY SIMILARITY.
FT DISULFID 450 487 BY SIMILARITY.
FT DISULFID 454 492 BY SIMILARITY.
FT DISULFID 465 477 BY SIMILARITY.
FT DISULFID 507 544 BY SIMILARITY.
FT DISULFID 511 549 BY SIMILARITY.
FT DISULFID 522 534 BY SIMILARITY.
FT DISULFID 554 565 BY SIMILARITY.
FT DISULFID 559 575 BY SIMILARITY.
FT DISULFID 578 589 BY SIMILARITY.
FT DISULFID 595 611 BY SIMILARITY.

```



```

FT DISULFID 602 BY SIMILARITY.
FT DISULFID 623 BY SIMILARITY.
FT DISULFID 647 BY SIMILARITY.
FT DISULFID 653 BY SIMILARITY.
FT DISULFID 660 BY SIMILARITY.
FT DISULFID 681 BY SIMILARITY.
FT DISULFID 708 BY SIMILARITY.
FT DISULFID 716 BY SIMILARITY.
FT DISULFID 721 BY SIMILARITY.
FT DISULFID 727 BY SIMILARITY.
FT DISULFID 751 BY SIMILARITY.
FT DISULFID 780 BY SIMILARITY.
FT DISULFID 800 BY SIMILARITY.
FT DISULFID 836 BY SIMILARITY.
FT DISULFID 839 BY SIMILARITY.
FT DISULFID 859 BY SIMILARITY.
FT DISULFID 877 BY SIMILARITY.
FT DISULFID 913 BY SIMILARITY.
FT DISULFID 949 BY SIMILARITY.
FT CARBOHYD 155 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 158 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 188 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 250 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 363 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 705 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 711 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1070 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1070 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1173 AA; 130019 MW; A9F036D516C0F24 CRC64;
SQ SEQUENCE 1173 AA; 130019 MW; A9F036D516C0F24 CRC64;

Query Match 5.6%; Score 266.5; DB 1; Length 1173;
Best Local Similarity 22.5%; Pred. No. 2.7e-12;
Matches 86; Conservative 54; Mismatches 147; Indels 95; Gaps 14;

QY 11 LLGIVLAWLRSQ-----AQSATVANVPVGPANPDLLPHLVEPEVDYIVKPKVLLVC- 65
DB 221 VEGTTLEAILRNKGLSMTNSVITLDNPVNGSPAIRTNVIGH-----KTKDLQAVCG 273
QY 66 -KAVPATOFFKNGEWRVQDVHVIETSTGSSGPTMEVRINVRQVQVEKVFGLLEYWC 124
DB 274 FSCDLSKLFAMKG-----LRTLVTUKDQVTKETKXNELIA 311
QY 125 QCVANSSGTTKSQKAYIRIARLRNFEQ-----EPLAKEVSLQEGIVLPC 170
DB 312 QIV-----TRTPGVCLHNGVLHKNRDEWTDSCTECTQNSATICKVSCF---LMPC 361
QY 171 RP---PEG-----IPPAEVEWLRNEDLVDPDLDPNVYITREHSLVVRQARLADTANY 219
DB 362 TNATIPDECCPCRPWPSADDDPWSWDTWTPCS-----VTCGHG-IQQRGRSCDSLNN 414
QY 220 TCVAKNIVAR-----RRSAAVIVVNGSGWTWTEWSVCASCGRGWQKRSRSCN 271
DB 415 PCGSSVQTRSCIQDCKRFQ-----DGNHSHWSPWSSCVTCGSGQITRLCNS 467
QY 272 PAPLNGGAFCEQNVHVRTVSSLLSVSDGSPWSKWACGLDC---THWRSRCSDDAP 328
DB 468 PVPQLNGKQCEGEGRENKPCQKPCPINGQWGPWSLMDTCTVTCGGGWQKRELRNNPKP 527
QY 329 RNGEFCOGTDLDRNCTSDLC 350
DB 528 QYEGKDCIGEPDTSQICNKKQDC 549

RESULT 9
SM5B_MOUSE STANDARD; PRT; 1093 AA.
AC Q60519;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Semaphorin 5B precursor (Semaphorin G) (Sema G).
GN SEMA5B OR SEMAG OR SEMG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10990;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI;
RX MEDLINE=96414430; PubMed=8817451;

RA Adams R.H., Betz H., Pueschel A.W.;
RT "A novel class of murine semaphorins with homology to thrombospondin
RL is differentially expressed during early embryogenesis.";
RL Mech. Dev. 57:33-45(1996).
CC -!- FUNCTION: May act as positive axonal guidance cues.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: In adult, only detected in brain.
CC -!- DEVELOPMENTAL STAGE: Differentially expressed in embryonic and
CC adult tissues. Its abundance decreases from E10 to birth.
CC -!- SIMILARITY: Belongs to the semaphorin family.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 7 TSP type-1 domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X97818; CAA66398.1; -.
DR MGD; MGI:107555; Sema5b.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP1.
DR Pfam; PF01437; PSI; 1.
DR Pfam; PF01403; Sema; 1.
DR Pfam; PF00090; tsp1; 5.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR SMART; SM00209; TSP1; 4.
DR SMART; PS50092; TSP1; 5.
DR PROSITE; PS50092; TSP1; 5.
DR Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 19
FT CHAIN 20 1093
FT DOMAIN 20 978 SEMAPHORIN 5B.
FT TRANSMEM 979 999 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 1000 1093 POTENTIAL.
FT DOMAIN 236 518 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 551 605 SEMA.
FT DOMAIN 606 662 TSP TYPE-1 1.
FT DOMAIN 664 713 TSP TYPE-1 2.
FT DOMAIN 721 776 TSP TYPE-1 3.
FT DOMAIN 795 850 TSP TYPE-1 4.
FT DOMAIN 852 907 TSP TYPE-1 5.
FT DOMAIN 908 952 TSP TYPE-1 6.
FT CARBOHYD 59 59 TSP TYPE-1 7.
FT CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 178 178 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 378 378 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 532 532 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 539 539 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 547 547 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 728 728 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 944 944 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1093 AA; 120326 MW; 29E5C9B1E8108717 CRC64;

Query Match 5.4%; Score 258.5; DB 1; Length 1093;
Best Local Similarity 22.2%; Pred. No. 1e-11;
Matches 104; Conservative 31; Mismatches 139; Indels 195; Gaps 16;

QY 163 EQGIVLPCRPEGIPPAEVEWLRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCV 222
DB 737 EQFRFTCRAP-----LPDP-----HGLQFGKRR---TETRTCP 767

```


DR PDB; ILSL; 18-DEC-02.
 DR GlycoSuiteDB; P07996; -.
 DR Genew; HGNC:11785; THBS1.
 DR MIM; 180600; -.
 DR GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.
 DR GO; GO:0004872; F:signal transducer activity; TAS.
 DR GO; GO:0007275; F:development; TAS.
 DR InterPro; IPR001881; EGF Ca.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR002110; IEGF.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP.
 DR InterPro; IPR003367; TSP_3.
 DR InterPro; IPR008859; TSPC.
 DR InterPro; IPR003129; TSPN.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00090; tsp_1; 3.
 DR Pfam; PF02412; tsp_3; 13.
 DR Pfam; PF05735; TSPC; 1.
 DR Pfam; PF02210; TSPN; 1.
 DR Pfam; PF00093; vwc; 1.
 DR PRINTS; PR01705; TSPIREPEAT.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00209; TSP1; 3.
 DR SMART; SM00210; TSPN; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS00022; EGF 1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00026; EGF_3; 2.
 DR PROSITE; PS00092; TSP1; 3.
 DR PROSITE; PS01208; WFC_1; 1.
 DR PROSITE; PS0184; WFC_2; 1.
 DR Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
 KW EGF-like domain; Signal; 3D-structure.
 FT SIGNAL 1 18
 FT CHAIN 19 1170
 FT DOMAIN 19 232
 FT DOMAIN 24 221
 FT DOMAIN 316 373
 FT DOMAIN 379 429
 FT DOMAIN 435 490
 FT DOMAIN 492 547
 FT DOMAIN 549 587
 FT DOMAIN 588 645
 FT DOMAIN 646 690
 FT DOMAIN 723 758
 FT DOMAIN 759 781
 FT DOMAIN 782 817
 FT DOMAIN 818 840
 FT DOMAIN 841 878
 FT DOMAIN 879 914
 FT DOMAIN 915 950
 FT DOMAIN 951 1170
 FT SITE 926 928
 FT DISULFID 270 270
 FT DISULFID 274 274
 FT DISULFID 391 423
 FT DISULFID 395 428
 FT DISULFID 406 413
 FT DISULFID 447 484
 FT DISULFID 451 489
 FT DISULFID 462 474
 FT DISULFID 504 541
 FT DISULFID 508 546
 FT DISULFID 519 531
 FT DISULFID 551 562
 FT DISULFID 556 572
 FT DISULFID 575 586
 FT DISULFID 592 608
 FT DISULFID 599 617
 FT DISULFID 620 644
 FT DISULFID 650 663

FT DISULFID 657 676 BY SIMILARITY.
 FT DISULFID 678 689 BY SIMILARITY.
 FT DISULFID 705 713 BY SIMILARITY.
 FT DISULFID 718 738 BY SIMILARITY.
 FT DISULFID 754 774 BY SIMILARITY.
 FT DISULFID 777 797 BY SIMILARITY.
 FT DISULFID 813 833 BY SIMILARITY.
 FT DISULFID 836 856 BY SIMILARITY.
 FT DISULFID 874 894 BY SIMILARITY.
 FT DISULFID 910 930 BY SIMILARITY.
 FT DISULFID 946 1167 BY SIMILARITY.
 FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 360 360 C-LINKED (MAN).
 FT CARBOHYD 385 385 /FTID-CAR 000205.
 FT CARBOHYD 394 394 /FTID-CAR 000206.
 FT CARBOHYD 438 438 C-LINKED (MAN).
 FT CARBOHYD 441 441 /FTID-CAR 000207.
 FT CARBOHYD 450 450 /FTID-CAR 000208.
 FT CARBOHYD 498 498 /FTID-CAR 000209.
 FT CARBOHYD 507 507 C-LINKED (MAN).
 FT CARBOHYD 507 507 /FTID-CAR 000210.
 FT CARBOHYD 507 507 O-LINKED (FUC. . .).
 FT CARBOHYD 708 708 /FTID-CAR 000211.
 FT CARBOHYD 1067 1067 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1067 1067 N-LINKED (GLCNAC. . .) (POTENTIAL).
 Query Match 5.1%; Score 246; DB 1; Length 1170;
 Best Local Similarity 30.3%; Pred. No. 1.1e-10;
 Matches 47; Conservative 24; Mismatches 66; Indels 18; Gaps 3;
 QY 207 VYRQARLADTANYTCVAKNIVAR-----RRSASAAVIVYVNGWSTWTSVCSASC 258
 Db 399 IQQGRGSCDNLNRCGSSVQRTCHIQCCKRFXQ-----DGGWHSWSPWSSCSVTC 451
 QY 259 GRGWKRSRSCINPAPLNGCAFCEGQNVHDRVSSLLVSDGWSWSPWSSCSAGLDLC--- 315
 Db 452 GDGVITRILCNPSQPMNGKFCGEARETKACKKDACPFNGWGWSPWSPWDCVTCGGG 511
 QY 316 THWRGRCSDPAPRNGGECQGTDLDTNRCTSDLC 350
 Db 512 VOKRSRLCNPPTQFGGKDCVGDVTENQICNKQDC 546
 RESULT 14
 TSP2_CHICK STANDARD; PRT; 1178 AA.
 ID TSP2_CHICK
 AC P35440;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Thrombospondin 2 precursor.
 GN THBS2 OR TSP2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91217026; PubMed=2022631;
 RA Lawler J., Duquette M., Ferro P.;
 RT "Cloning and sequencing of chicken thrombospondin."
 RL J. Biol. Chem. 266:8039-8043(1991).
 CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
 cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
 laminin and type V collagen.
 CC -!- SUBUNIT: Homotrimer; disulfide-linked.
 CC -!- SIMILARITY: Belongs to the thrombospondin family.
 CC -!- SIMILARITY: Contains 1 VWF domain.

FT	DISULFID	559	570	BY SIMILARITY.
FT	DISULFID	564	580	BY SIMILARITY.
FT	DISULFID	583	594	BY SIMILARITY.
FT	DISULFID	600	616	BY SIMILARITY.
ET	DISULFID	607	625	BY SIMILARITY.
FT	DISULFID	628	652	BY SIMILARITY.
FT	DISULFID	658	671	BY SIMILARITY.
FT	DISULFID	665	684	BY SIMILARITY.
FT	DISULFID	686	697	BY SIMILARITY.
FT	DISULFID	713	721	BY SIMILARITY.
ET	DISULFID	726	746	BY SIMILARITY.
FT	DISULFID	762	782	BY SIMILARITY.
FT	DISULFID	785	805	BY SIMILARITY.
FT	DISULFID	821	841	BY SIMILARITY.
FT	DISULFID	844	864	BY SIMILARITY.
FT	DISULFID	882	902	BY SIMILARITY.
FT	DISULFID	918	938	BY SIMILARITY.
FT	DISULFID	954	1175	BY SIMILARITY.
FT	CARBOHYD	157	157	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	244	244	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	317	317	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	322	322	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	463	463	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	590	590	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	716	716	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1075	1075	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE	1178 AA,	131816 MW, F37E02F42C8717A2 CRC64,	(POTENTIAL).

Query Match 5.1%; Score 243.5; DB 1; Length 1178;
Best Local Similarity 34.4%; Pred. No. 1.7e-10;
Matches 55; Conservative 15; Mismatches 75; Indels 15; Gaps 4;

QY	210 QARLATANTYCAKNIVARRRS-ASAAVIVYVGWSTWTWSVCSSACGRGWKSR	268
Db	410 RGRSCDVTRSACTGPHIQTRMCSFKKHRIODGGWHSWPSSCSVTGGVNITRIL	469
QY	269 CTNPAPLNGAFCEGQNVHVRTVSLLSVSDGSWSPSKWSACGLDC---THWRSECD	325
Db	470 CNSPIQMGGKNCVGNGRETEKECAPCPVNGQGPWSPNSACTVTCGGIRERSLCNS	529
QY	326 PAPERNGEBEQGTDLDT-----RNCTSDLCVHSASGP	357
Db	530 PEPOYGKPCVG--DTKQHDMCKRDCPIDGCLSNCPF	566

RESULT 15
TSP1_BOVIN STANDARD; PRT; 1170 AA.
ID TSP1_BOVIN STANDARD; PRT; 1170 AA.
AC Q28178; Q28179;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Thrombospondin 1 precursor.
GN THBS1 OR TSP1 OR TSP-1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Holstein; TISSUE-Tooth;
RX MEDLINE=98173773; PubMed=9507054;
RA Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S., Inoue H.;
RA "cDNA cloning of bovine thrombospondin 1 and its expression in odontoblasts and preodontoblasts."
RL Biochim. Biophys. Acta 1382:17-22(1998).
RN [2]
RN SEQUENCE OF 1-18 AND 710-1170 FROM N.A.
RP TISSUE=Aortic endothelium;
RC Zafar R.S., Moil Y.D., Womack J.F., Walz D.A.;
RA Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases

Search completed: October 6, 2004, 18:18:00
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 17:50:33 ; Search time 84 Seconds

(without alignments)
3376.796 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRPGIWPALLGIVLAWL.....AVAGLQPDAGLFTVSEAE 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	4592.5	95.9	898	11	Q8K1S4	Q8K1S4 mus musculus
2	4545.5	95.0	898	11	O08721	O08721 rattus norv
3	2819.5	58.9	544	4	Q96GP4	Q96GP4 homo sapien
4	2705.5	56.5	931	11	O08747	O08747 mus musculus
5	2686	56.1	950	11	O8CD16	O8CD16 mus musculus
6	2679.5	56.0	931	13	Q7T2Z5	Q7T2Z5 gallus gall
7	2673.5	55.8	931	4	O95185	O95185 homo sapien
8	2570.5	53.7	943	13	Q8UGT4	Q8UGT4 xenopus lae
9	2505.5	52.3	1008	11	O80Y85	O80Y85 mus musculus
10	2499	52.2	945	11	O081S3	O081S3 mus musculus
11	2497	52.2	945	11	O08722	O08722 rattus norv
12	2493	52.1	945	11	Q9D398	Q9D398 mus musculus
13	2486.5	51.9	934	4	O81ZJ1	O81ZJ1 homo sapien
14	2479	51.8	945	4	Q86SN3	Q86SN3 homo sapien
15	2144.5	44.8	956	11	Q8K1S2	Q8K1S2 mus musculus
16	2136	44.6	948	4	Q8WVP7	Q8WVP7 homo sapien

17	1597.5	33.4	597	4	Q8IU70	Q8IU70 homo sapien
18	1442	30.1	328	11	Q80T71	Q80T71 mus musculus
19	1233.5	25.6	554	4	Q8NIY2	Q8NIY2 homo sapien
20	971	20.3	365	4	Q9H9F3	Q9H9F3 homo sapien
21	961.5	20.1	1072	5	Q9NBL0	Q9NBL0 drosophila
22	956.5	20.0	1072	5	Q9V7B5	Q9V7B5 drosophila
23	953.5	19.9	947	5	Q26262	Q26262 caenorhabdi
24	950.5	19.9	947	5	O44171	O44171 caenorhabdi
25	694	14.5	199	13	Q9PVD5	Q9PVD5 petromyzon
26	552.5	11.5	351	4	Q8TF26	Q8TF26 homo sapien
27	366	7.6	2673	4	Q96SC3	Q96SC3 homo sapien
28	366	7.6	5636	4	Q96RW7	Q96RW7 homo sapien
29	311.5	6.5	325	5	O81IK1	O81IK1 drosophila
30	306	6.4	518	4	O8IV45	O8IV45 homo sapien
31	273.5	5.7	1172	11	Q8CG21	Q8CG21 mus musculus
32	273.5	5.7	1172	11	Q7TMT3	Q7TMT3 mus musculus
33	271	5.7	1582	11	Q8CGM0	Q8CGM0 mus musculus
34	270.5	5.7	1081	5	Q9U631	Q9U631 drosophila
35	269.5	5.6	1083	5	Q9VTT0	Q9VTT0 drosophila
36	269.5	5.6	1091	5	Q7YU67	Q7YU67 drosophila
37	267	5.6	1522	11	Q8OZF8	Q8OZF8 mus musculus
38	263.5	5.5	1461	5	Q8MYA8	Q8MYA8 caenorhabdi
39	258.5	5.4	1122	11	Q7TT33	Q7TT33 mus musculus
40	258	5.4	1573	4	Q8NGW8	Q8NGW8 homo sapien
41	256	5.3	478	11	O8BVE5	O8BVE5 mus musculus
42	254.5	5.3	1560	11	Q8CGM1	Q8CGM1 mus musculus
43	249	5.2	685	6	Q9TTS5	Q9TTS5 bos taurus
44	249	5.2	5146	6	Q8SPM4	Q8SPM4 bos taurus
45	248	5.2	1171	11	Q8CGB2	Q8CGB2 mus musculus

ALIGNMENTS

RESULT 1

Q8K1S4	PRELIMINARY;	PRT;	898 AA.
ID	Q8K1S4		
AC	Q8K1S4;		
DT	01-OCT-2002 (Tremblrel. 22, Created)		
DT	01-OCT-2002 (Tremblrel. 22, Last sequence update)		
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)		
DE	Netrin receptor Unc5h1.		
GN	UNC5H1.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Engelkamp D.;		
RT	"Cloning of three mouse unc-5 genes and their expression patterns at mid-gestation."		
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AJ487852; CAD32250.1; --		
DR	MGI; 834682; Unc5h1.		
DR	GO; GO:0004872; P:receptor activity; IEA.		
DR	GO; GO:0007165; P:signal transduction; IEA.		
DR	InterPro; IPR000488; Death.		
DR	InterPro; IPR003599; IG.		
DR	InterPro; IPR007110; IG-like.		
DR	InterPro; IPR000884; TSP1.		
DR	InterPro; IPR008085; TSP1.		
DR	InterPro; IPR000906; ZUS.		
DR	Pfam; PF00531; death; 1.		
DR	Pfam; PF00047; ig; 1.		
DR	Pfam; PF00090; tsp_1; 2.		
DR	Pfam; PF00791; ZUS_1; 1.		
DR	PRINTS; PR01705; TSP1REPEAT.		
DR	SMART; SM00005; DEATH; 1.		
DR	SMART; SM00409; IG; 1.		
DR	SMART; SM02403; TSP1; 2.		
DR	SMART; SM00218; ZUS; 1.		
DR	PROSITE; PS50835; IG_LIKE; 1.		


```

Db 420 LTIQPDLSSTTTTQGSCLCPQDGPSPKQFQLSNGHLLSPLGGRHTLHSSPTSEAEFV 479
QY 480 SRLSTQNYFRSLPRGTSNMTYGTENFLGRLMIPNTGISLLIPDAIPRGKIYEIYLT 539
Db 480 SRLSTQNYFRSLPRGTSNMTYGTENFLGRLMIPNTGISLLIPDAIPRGKIYEIYLT 539
QY 540 KPEDVRLPLAGCQTLSPISVSCGPPGVLLTRPVILAMDHCGEPPSPDSWSLRKQSCGS 599
Db 540 KPEDVRLPLAGCQTLSPISVSCGPPGVLLTRPVILAMDHCGEPPSPDSWSLRKQSCGS 599
QY 600 WEQDVLHGEAPSHLYYCQLEASACVYFTQGRFALVGEALSVAALKRLKLLPAPVA 659
Db 600 WE-DVLHGEAPSHLYYCQLEASACVYFTQGRFALVGEALSVAALKRLKLLPAPVA 658
QY 660 CTSLEYNIRVYCLDHTDALKEVVOLEKQGGQIQEPRVLHFKDSYHNLRSLSHDVPSS 719
Db 659 CTSLEYNIRVYCLDHTDALKEVVOLEKQGGQIQEPRVLHFKDSYHNLRSLSHDVPSS 718
QY 720 LWSKLLVSYQEIIPFYHIWNGTQRYLHCTFTLERSVSPSTSDACKLWVWQVGGQGSFSI 779
Db 719 LWSKLLVSYQEIIPFYHIWNGTQRYLHCTFTLERSVSPSTSDACKLWVWQVGGQGSFNI 778
QY 780 NFNITKDTREAEALLALESEAGVPALVGPSPAFKIPFLIRQKIISSLDPPCRRGADWRTLQ 839
Db 779 NFNITKDTREAEALLALESEAGVPALVGPSPAFKIPFLIRQKIISSLDPPCRRGADWRTLQ 838
QY 840 KLHLDHSLSPFASKPSPMTAMILNWEARHPNGNLSQAAVAGLGOPDAGLFTVSEAE 899
Db 839 KLHLDHSLSPFASKPSPMTAMILNWEARHPNGNLSQAAVAGLGOPDAGLFTVSEAE 898

RESULT 3
Q96GP4
ID Q96GP4 PRELIMINARY; PRT; 544 AA.
AC Q96GP4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similar to transmembrane receptor Unc5H1 (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC009333; AAH09333.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZUS; 1.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZUS; 1.
KW Receptor; Transmembrane.
FT NON_TER 1
SQ SEQUENCE 544 AA; 5949 aa; 350A7BA53375CCAE CRC64;

Query Match 58.9%; Score 2819; DB 4; Length 544;
Best Local Similarity 99.6%; Pred. No. 8.2e-252;
Matches 540; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 359 DVALYVGLIAVAVCLVLLVLLVLYVCRKKGLEDSDVADSSILTSQFQVSVTKPSKADNP 418
Db 4 DVALYVGLIAVAVCLVLLVLLVLYVCRKKGLEDSDVADSSILTSQFQVSVTKPSKADNP 63
QY 419 HLLTTPDLS-TTTTYQSLCPQDGPSPKQFQLTNGHLLSPLGGRHTLHSSPTSEAE 477

```

```

Db 64 HLLTTPDLS-TTTTYQSLCPQDGPSPKQFQLTNGHLLSPLGGRHTLHSSPTSEAE 123
QY 478 FVSRLSTQNTFRSLPRGTSNMTYGTENFLGRLMIPNTGISLLIPDAIPRGKIYEIYLT 537
Db 124 FVSRLSTQNTFRSLPRGTSNMTYGTENFLGRLMIPNTGISLLIPDAIPRGKIYEIYLT 183
QY 538 LHKPEDVRLPLAGCQTLSPISVSCGPPGVLLTRPVILAMDHCGEPPSPDSWSLRKQSC 597
Db 184 LHKPEDVRLPLAGCQTLSPISVSCGPPGVLLTRPVILAMDHCGEPPSPDSWSLRKQSC 243
QY 598 GSWEQDVLHGEAPSHLYYCQLEASACVYFTQGRFALVGEALSVAALKRLKLLPAP 657
Db 244 GSWE-DVLHGEAPSHLYYCQLEASACVYFTQGRFALVGEALSVAALKRLKLLPAP 302
QY 658 VACTSLEYNIRVYCLDHTDALKEVVOLEKQGGQIQEPRVLHFKDSYHNLRSLSHDVP 717
Db 303 VACTSLEYNIRVYCLDHTDALKEVVOLEKQGGQIQEPRVLHFKDSYHNLRSLSHDVP 362
QY 718 SSWKSKLLVSYQEIIPFYHIWNGTQRYLHCTFTLERSVSPSTSDACKLWVWQVGGQGSF 777
Db 363 SSWKSKLLVSYQEIIPFYHIWNGTQRYLHCTFTLERSVSPSTSDACKLWVWQVGGQGSF 422
QY 778 SINFNITKDTREAEALLALESEAGVPALVGPSPAFKIPFLIRQKIISSLDPPCRRGADWRTL 837
Db 423 SINFNITKDTREAEALLALESEAGVPALVGPSPAFKIPFLIRQKIISSLDPPCRRGADWRTL 482
QY 838 AQKLHLDHSLSPFASKPSPMTAMILNWEARHPNGNLSQAAVAGLGOPDAGLFTVSEA 897
Db 483 AQKLHLDHSLSPFASKPSPMTAMILNWEARHPNGNLSQAAVAGLGOPDAGLFTVSEA 542
QY 898 EC 899
Db 543 EC 544

RESULT 4
O08747
ID O08747 PRELIMINARY; PRT; 931 AA.
AC O08747;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Rostral cerebellar malformation protein.
GN UNC5H3 OR RCM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57B6/SJL;
RX MEDLINE=97271898; PubMed=9126743;
RA Ackerman S.L., Kozak L.P., Przyborski S.A., Rund L.A., Boyer B.B.,
RA Knowles B.B.;
RT "The mouse rostral cerebellar malformation gene encodes an UNC-5-like
protein."
RL Nature 386:838-842(1997).
DR EMBL; U72634; AAB54103.1; -.
DR MGD; MGI:1095412; Unc5h3.
DR GO; GO:0005886; C:plasma membrane; IC.
DR GO; GO:0005042; F:netrin receptor activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0007420; F:brain development; IMP.
DR GO; GO:0030334; P:regulation of cell migration; IMP.
DR InterPro; IPR000488; Death.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP1.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; tsp_1; 2.

```

DR	Pfam; PF00791; ZUS; 1.	858	APQTRGDWRMLAHKLNLDRLNYFATKSSPTGVILDWEAQNFDPGNLSMLAAVLEWVG	917
DR	PRINTS; PR01705; TSP1REPEAT.			
DR	SMART; SM00005; DEATH; 1.	886	QPDAGLFTVSEAE	898
DR	SMART; SM00408; IGC2; 1.		:	:
DR	SMART; SM00209; TSP1; 2.	918	RHETVVSAAEQG	930
DR	SMART; SM00218; ZUS; 1.		:	:
DR	PROSITE; PS00835; IG LIKE; 1.		:	:
DR	PROSITE; PS50092; TSP1; 2.		:	:
KW	Immunoglobulin domain.			
SQ	SEQUENCE 931 AA; 103062 MW; 8A5D951A4BECA179 CRC64;			
Query Match 56.5%; Score 2705.5; DB 11; Length 931;				
Best Local Similarity 56.4%; Pred. No. 6.4e-241;				
Matches 515; Conservative 153; Mismatches 214; Indels 31; Gaps 12;				
QY	9	PALLGVLAALRGSGAQQA---TVANPVPGANPDLLPHFLVEPEDVVIYVKNKPVLLVC	65	
DB	26	PAL--ALLSAGTGSAAQDDFFHELPEFPDPPELPHFLIEPEAYIVKNKPVNLYC	83	
QY	66	KAVPATQIIPKCNSEWVRQVDHVIERTSDGSSGEPTEWVRINVSROQVEKVGLEBYWCQ	125	
DB	84	KASPATQIYFKCNSEWVHQDHVDVDERVDETSGLIVREVSIEISROQVEELFGPDYWCQ	143	
QY	126	CVANSSGTTTQSKAYIRIARLKNPEOEPLAKEVSLQGIIVLPCRPEGIPPAEVEWLR	185	
DB	144	CVANSSAGTTKSKAYIRIARLKNPEOEPLAKEVSLQGIIVLPCRPEGIPPAEVEWLR	185	
QY	826	PPCRGADWRTLAOKLHDSLSFFASKPSPTAMILNLWEARHFPNGNLSQALAAVAGLG	885	
Query Match 56.1%; Score 2686; DB 11; Length 950;				
Best Local Similarity 55.3%; Pred. No. 4.1e-239;				
Matches 515; Conservative 153; Mismatches 214; Indels 50; Gaps 13;				
QY	9	PALLGVLAALRGSGAQQA---TVANPVPGANPDLLPHFLVEPEDVVIYVKNKPVLLVC	65	
DB	26	PAL--ALLSAGTGSAAQDDFFHELPEFPDPPELPHFLIEPEAYIVKNKPVNLYC	83	
QY	66	KAVPATQIIPKCNSEWVRQVDHVIERTSDGSSGEPTEWVRINVSROQVEKVGLEBYWCQ	125	
DB	84	KASPATQIYFKCNSEWVHQDHVDVDERVDETSGLIVREVSIEISROQVEELFGPDYWCQ	143	
QY	126	CVANSSGTTTQSKAYIRIARLKNPEOEPLAKEVSLQGIIVLPCRPEGIPPAEVEWLR	185	
DB	144	CVANSSAGTTKSKAYIRIARLKNPEOEPLAKEVSLQGIIVLPCRPEGIPPAEVEWLR	203	

DR	EMBL; AY187310; AAC067275.1; -.																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	</
----	---------------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

O95185	PRELIMINARY;	PRT;	931 AA.
ID	O95185;		
AC	O95185;		
DT	01-WAY-1999 (TtEMBLrel. 10, Created)		
DT	01-WAY-1999 (TtEMBLrel. 10, Last sequence update)		
DT	01-OCT-2003 (TtEMBLrel. 25, Last annotation update)		
DE	Transmembrane receptor UNC5C.		
UN	UNC5C.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Brain;		
RX	MEDLINE=99000841; PubMed=9782087;		
RA	Ackerman S.L., Knowles B.B.;		
RT	"Cloning and mapping of the UNC5C gene to human chromosome 4q21-q23.;"		
RL	Genomics 52:205-208 (1998).		
DR	EMBL; AF055634; AAC67491.1; -.		
DR	Genew; HGNC:12569; UNC5C.		
DR	GO; GO:0005042; Finitrin receptor activity; TAS.		
DR	GO; GO:0007411; P:axon guidance; TAS.		
DR	GO; GO:0007420; P:brain development; TAS.		
DR	InterPro; IPR000488; Death.		
DR	InterPro; IPR007110; Ig-like.		
DR	InterPro; IPR003598; Ig c2.		
DR	InterPro; IPR000884; TSP1.		
DR	InterPro; IPR008085; TSP 1.		
DR	InterPro; IPR000906; ZUS.		
DR	Pfam; PF00531; death; 1.		
DR	Pfam; PF00047; Ig; 1.		
DR	Pfam; PF00090; tsp_1; 2.		
DR	Pfam; PF00791; ZUS; 1.		
DR	PRINTS; PR01705; TSP1REPEAT.		
DR	SMART; SM00005; DEATH; 1.		
DR	SMART; SM00408; IGC2; 1.		
DR	SMART; SM00209; TSP1; 2.		
DR	SMART; SM00218; ZUS; 1.		
DR	PROSITE; PS50835; IG_LIKE; 1.		
DR	PROSITE; PS50092; TSP1; 2.		
KW	Immunoglobulin domain; Receptor.		
SQ	SEQUENCE 931 AA; 103101 MW, EFD71122C98DABB8 CRC64;		
Query Match 55.8%; Score 2673.5; DB 4; Length 931;			
Best Local Similarity 55.5%; Pred. No. 5.7e-238;			
Matches 507; Conservative 154; Mismatches 221; Indels 31; Gaps 12;			
Qy	9	PALLGIVLAALRGSGAQQS---ATVANPVGANPDLLPHFLVPEPDVYVKNKPVLVC	65
Db	26	PAL--ALLASGTGSAQQDDFFHELPEPFPDPEPLPHFLIEPEAYIVKNKPNVLYC	83
Qy	66	KAVPATQIFPKCGEWEVRQVDHVIERSDGGSGEPTMEVRINVSRRQVEKVFGLSEYWCQ	125
Db	84	KASPATQIFKCNSEWHQKHDIHVDERSGLIVREVSIEISRQVEELFGEDYWCQ	143
Qy	126	CVWSSSGTTKSKAYIRIARLKNFQEPLEAKVSLFQGLVLCRPEPGIPPAEVEWLR	185
Db	144	CVWSSAGTTKSRAYIRIARLKNFQEPLEAKVSLFQGLVLCRPEPGIPPAEVEWLR	203
Qy	186	NEDLVDPSPNVYITREHSIVVQARLADTANYTCVAKNIVARRRSASAAVIVVNGW	245
Db	204	NEDIIDVEDRNFYITIDHNLIIQARLSDTANTVCVAKNIVARRRSATVIVVNGW	263
Qy	246	STWTWSSYCSACGRGQKRSCTNPAPLNGGAFCEGQNVHRTVSSLLVSDGSPW	305
Db	264	STWTWSSYCSACGRGQKRSCTNPAPLNGGAFCEGQSV-QKTACTTLCFVDGRTPW	322
Qy	306	SKWSACGLDCTHWSRSCDPAFNGGEECGTDLDTFRNCTSDLCVHSASGPDVLYVG	365
Db	323	SKWSTCGTECHWRRECTAPAKNGGKDCDGLVLQSKNCTDGLCMQTPADSDVLYVG	382
Qy	366	L-IAVAVCLVLLVLLVILVCRKKEGLSDVDADSSILTSFGFPVSIKPSKADNPHLLTIQ	424

Db	383	IVIAIVIVCLAIISVVVALFYRKNHRDFESDIIDSSALNGFGFQPVNIIKAARQD---LLAVP	439
Qy	425	PDL-STTTTYQGSCLPRQGPSPKFLQTNHLLSLPLGGGRHLHLS-----SPTSEAEFV	479
Db	440	PDLTSAAMYRGFVVALHD-VSDKIPMTNSPILDLPLNLKI KYNTSGAVSPQDDISEFT	498
Qy	480	SRLS---TQNYF-----RSLPRGT--SNMTYGTENFLGRGLMIPMTGILLIPDPA	525
Db	499	SKLSPOMTOSLLENEALSUKNSLARQTPDSCFAGSFLNSLGGHLIVFNSGVSLIPAGA	558
Qy	526	IPRGKIYEIYLTILHKPEDVRLPLAGCQTLTSLPIVSCGPPGVLLTRPVILLAMDHCGEPPSD	585
Db	559	IPQGRVYEMVTVHRKETMRPPMDDSQTLTTPVSCGPPGALLTRPVLTMRHCADPNT	618
Qy	586	SWSLRLKKQSCGSEQDVLHGEAPSHLYYCOLEASACYVTEQIGRFPALVGEALSVA	645
Db	619	DKWILLKNOAAQOWE-DVVVVGEENFTPCYIKLDAEACHILTENLSTYALVGHSTTKA	677
Qy	646	AAKRLKLLPAPVACTSLEVNIHVYCLHDTHDALKWVLEKQLGQOLIQEPRVLHPKDS	705
Db	678	AAKRLKLLPAPVACTSLEVNIHVYCLHDTHDALKWVLEKQLGQOLIQEPRVLHPKDS	737
Qy	706	YHNLRLSIHNDVPSLWKLVSQYIPIFYHWNQYRVLHCTFTTLERVSPTSDDLAKL	765
Db	738	THNLRLSIHNDVPSLWKLVSQYIPIFYHWNQYRVLHCTFTTLERVSPTSDDLAKL	797
Qy	766	WVWQVGGGQSPSINFTKDFPAELLALLESAGVPALVGPSPAFKIPFLIROLKLSLD	825
Db	798	CVRQVEGEGQIFQINCTVSEETGIDPLLDPAITTTVTGSPAFSIPFLIROLKLSLD	857
Qy	826	PPCERGADETLAKLHLSHLSFFASKPSPTAMILNLEWHPFNGLSOLAAAVAGLG	885
Db	858	APQTRGHDMRLAKNLDRILNFATKSPGTGVLIDLEWAEQFPGNLSLAAVLEMG	917
Qy	886	QPDAGLFTVSEAE 898	
Db	918	RHETVVSLLAEQG 930	
RESULT 8			
Q8JGT4 PRELIMINARY; PRT; 943 AA.			
ID	Q8JGT4		
AC	Q8JGT4		
DT	01-OCT-2002 (TtEMBLrel. 22, Created)		
DT	01-OCT-2002 (TtEMBLrel. 22, Last sequence update)		
DT	01-OCT-2003 (TtEMBLrel. 25, Last annotation update)		
DE	UNC-5 receptor.		
OS	Xenopus laevis (African clawed frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;		
OC	Xenopodinae; Xenopus.		
OX	NCBI_TaxID=8355;		
RN	[1]		
RA	SEQUENCE FROM N.A.		
RA	Anderson R.B., Holt C.E.;		
RT	"Expression of UNC-5 in the developing Xenopus visual system.;"		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY099459; AAM34486.1; -.		
DR	GO; GO:0004872; F:receptor activity; IEA.		
DR	GO; GO:0007165; P:signal transduction; IEA.		
DR	InterPro; IPR000488; Death.		
DR	InterPro; IPR003599; Ig.		
DR	InterPro; IPR007110; Ig-like.		
DR	InterPro; IPR003598; Ig c2.		
DR	InterPro; IPR00884; TSP1.		
DR	InterPro; IPR008085; TSP 1.		
DR	InterPro; IPR000906; ZUS.		
DR	Pfam; PF00531; death; 1.		
DR	Pfam; PF00047; Ig; 1.		
DR	Pfam; PF00090; tsp_1; 2.		
DR	Pfam; PF00791; ZUS; 1.		
DR	PRINTS; PR01705; TSP1REPEAT.		

QY 401 LTSGFQFVSIKPKADNPHELL--TIQPDLSITT--TYQGSICPRQDGPSPKQFQTLNGHLLS 457
 Db 417 LTGGEFVNFKTARPNNPQLLHPASPDLTASAGIRGVPYALQDS--ADKIPMTNPSPLLD 475
 QY 458 PLGGGRHTLHSS-----PTSEAEFVSRLSTQNYFRS----- 490
 Db 476 PLPSLKITVYNSSTIGSGGLADGADLLGLVPPGTYPGDF--SRDTHFLHLSASLGSQHL 534
 QY 491 --LPRGTNNMTYGTENFLGRLMIPNTGISLLIPDAIPRGKIYEIYTLHKEDVRLPL 548
 Db 535 LGLPRDPSSVSGTFCGLGRLSLPGTGVSLVFNPAIGPKGYDGLYHINKAEST--LPL 593
 QY 549 A--GCQTLLSPIVSCGPPGVLLTPVILAMDHCHCEPSPDSLSRLKKQSCGSEQDVLHL 607
 Db 594 SEGQTVLSPSVTCGTGLLCPVLLTPVPHCAEVTAAGDWIFOLKTOAHGQHW--EVVTL 652
 QY 608 GEAPSHLYYQLEASACVYFTQLGRFALVGEALSVAAAKRLKLLFAPVACTSLEYNI 667
 Db 653 DEETLNTPCYQLEAKSCHILLDQGTGVFMGESYSRSVAKRLQLAIFAPALCTSLSEYSL 712
 QY 668 RYVCLDTHDALKEVVOLEKQGLIQPRVLHFKDSYHNRLSLDHPSSLWKSLLV 727
 Db 713 RYVCLDTPVALKEVLELERTLGGYLVBEPKPLLFKDSYHNRLSLDHPHAWRSKLLA 772
 QY 728 SYQEIPIFYHNGTQRYLHCTFTLERSVSPSTSLACKLWVWQVGGQGSINFINITKDT 787
 Db 773 KYQEIPIFYHNGSQRALHCTFTLERSHSLASTFTCKVQVQVGGQGFQHLHTTIA--ET 831
 QY 788 RFAELLALSEAG--VPALVGSFAKIPFLIROKIISSLDPPCRRGADWRTLAQKHLDS 845
 Db 832 PAGESLDALCSAPGNAITQGLPYAFKIPLSIROKICSLDAPNSRGNDWELLAAQLSMDR 891
 QY 846 HLSFFASKPSPMTAILNLEAFHPNGLSOLAAVAGIQPDAGLFTVSEARC 899
 Db 892 YLNYFATKASPTGVILDLWEARQQDDGLNSLASALEENGKSEMLVAMATDGDG 945

RESULT 11
 O08722 PRELIMINARY; PRT; 945 AA.

AC O08722;
 DT 01-JUL-1997 (trEMBLrel. 04, Created)
 DT 01-JUL-1997 (trEMBLrel. 04, Last sequence update)
 DT 01-OCT-2003 (trEMBLrel. 25, Last annotation update)
 DE Transmembrane receptor UNC5H2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN NCBI_TaxID=10116;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97271897; PubMed=9126742;
 RA Leonardo E.D., Hincin L., Masu M., Keino-Masu K., Ackerman S.L.,
 RA Tessier-Lavigne M.;
 RT "Vertebrate homologues of C. elegans UNC-5 are candidate netrin
 RT receptors";
 RL Nature 386:833-838 (1997).
 DR EMBL; U87306; AAB57679.1; -.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR000885; TSP_1.
 DR InterPro; IPR000906; ZUS.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; tsg_1; 2.
 DR Pfam; PF00791; ZUS; 1.
 DR PRINTS; PRO1705; TSP1REPEAT.
 DR SMART; SM00005; DEATH; 1.

DR SMART; SM00408; IGC2; 1.
 DR SMART; SM00209; TSP1; 2.
 DR SMART; SM00218; ZUS; 1.
 DR PROSITE; PS00017; DEATH_DOMAIN; 1.
 DR PROSITE; PS00835; IG_LIKE; 1.
 DR PROSITE; PS00092; TSP1; 2.
 KW Immunoglobulin domain; Receptor.
 SQ SEQUENCE 945 AA; 103520 MW; 6B9C2A262E560B9B CRC64;

Query Match 52.2%; Score 2497; DB 11; Length 945;
 Best local similarity 52.0%; Pred. No. 1.3e-221;
 Matches 501; Conservative 144; Mismatches 236; Indels 82; Gaps 20;

QY 1 MAVRGLWALLCIVLAAM-----LRG--SQAQOSTAVNPVPCANPDLLPHELVPEDV 53
 Db 1 MRARSARGALLALLLCLWDPTPSLAGIDSGGQ---ALPDSFSAFAEQLPHELLEPEDA 57
 QY 54 YIVKKNPVLIVCKAVPATQIFPKCNGEWRQVDHVIERTSDGSSGPTMEVRINVSRQV 113
 Db 58 YIVKKNPVELHCRAPFATQIYFKCNGEWSQKHVTQESLDEATGURIREVQIEVSRQV 117
 QY 114 EKVFGLSEYWCQCAWSSSGTTKSKAYIRIARLRNFEQELAKEVSELEQGIIVLCRPP 173
 Db 118 EELFGLSEYWCQCAWSSSGTTKSRAYIRIAYLRKNFQELAKEVPLDHEVLLQCRPP 177
 QY 174 EGTPPAEVEWLNEDLVPSLDPNVYITREHSLVVRQALADTANTYCAKNIIVARRSA 233
 Db 178 EGVFAEVEWLNEDVIDPAQDTNFLTIDHNLIRQARLSDTANTYCAKNIIVARRST 237
 QY 234 SAAVIVYVNGWSTWTEWSVCASCGRWQKRSRCTNPAPLNGGAFCEGQNVHSDRTVSS 293
 Db 238 TATVIVYVNGWSSWAEWSPCNRGCRGWQKTRTCTINPAPLNGGAFCEGQ--ACQIATCT 296
 QY 294 LIVSDGWSWMSKWSACGLDCTHMSRECSPPAPENGEECGTDLDTNCTSDLCV-- 351
 Db 297 TVCPVDGAWTEWSKWSACSTECAHWSRECMAPPQNGGRCDSGLTDSKNCCTDGLCVLN 356
 QY 352 --HSASGE-----DVALYVGL--IAVAVCLVLLLVILVYCRKKEGLSDVAOSS--I 400
 Db 357 QRTLNDPKSRPLEPSGDVALYAGLVAVFVVLAVLAVMAGVIVYRRNCRDFDITDSSAA 416
 QY 401 LTSGFQFVSIKPKADNPHELL--TIQPDLSITT--TYQGSICPRQDGPSPKQFQTLNGHLLS 457
 Db 417 LTGGEFVNFKTARPNNPQLLHPASPDLTASAGIRGVPYALQDS--ADKIPMTNPSPLLD 475
 QY 458 PL-----GGG-----RHTLHSSPTSEAEFV 479
 Db 476 PLPSLKITVYNSSTIGSGGLADGADLLGLVPPGTYPGDFSRDTHFLHLS----- 526
 QY 480 SRLSTQNYFRSLPRGTNNMTYGTENFLGRLMIPNTGISLLIPDAIPRGKIYEIYTLH 539
 Db 527 ASLGSQ--HLLGLPRDPSSVSGTFCGLGRLTIPGTGVSLVFNPAIGPKGYDGLYLRIN 585
 QY 540 KPEDVRLPLA--GCQTLLSPIVSCGPPGVLLTPVILAMDHCHCEPSPDSLSRLKKQSCG 598
 Db 586 KTEST--LPLSEGSQVILSPSVTCGTGLLCPVLLTPVPHCAEVTAAGDWIFOLKTOAHG 644
 QY 599 SWEQDVLHGEAPSHLYYQLEASACVYFTQLGRFALVGEALSVAAAKRLKLLFAPV 658
 Db 645 HWE--EVVTLDEBTLNTPCYQLEAKSCHILLDQGTGVFMGESYSRSVAKRLQLAIFAPA 703
 QY 659 ACTSLEYNIYRVCYLDHDAKKEVVOLEKQGLIQPRVLHFKDSYHNRLSLDHPSS 718
 Db 704 LCTSLEYSLRYVCLDTPVALKEVLELERTLGGYLVBEPKPLLFKDSYHNRLSLDHP 763
 QY 719 SLWKSLLVSYQEIPIFYHNGTQRYLHCTFTLERSVSPSTSLACKLWVWQVGGQGS 778
 Db 764 AHWSKLLAKYQEIPIFYHNGSQRALHCTFTLERSHSLASTFTCKVQVQVGGQGFQ 823
 QY 779 INFNITKDTFRFAELLALSEAGVPAI--VGPSAFKIPFLIROKIISSLDPPCRRGADWRT 836
 Db 824 LHTTLA--ETPAGSLDALCSAPGNAITQGLPYAFKIPLSIROKICSLDAPNSRGNDWEL 882

QY 837 LAQKLDHSHLFFASKSPPTAMINLWEARHPNGLNSQLAAAVAGLQPDAGLFTVSE 896
 DB 883 LAQKLSMDRYLNFATKASPTGVLIDWEARQDDGLNSLASALEEMKSEMLVAMTTD 942

QY 897 AEC 899
 DB 943 GDC 945

RESULT 12

Q9D398 PRELIMINARY; PRT; 945 AA.
 AC Q9D398;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE 6330415802Rik protein.
 GN UNC5H2 OR 6330415802RIK.
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK018177; BAB31108.1; -.
 DR MGD; MGI:894703; Unc5h2.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig c2.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP1.
 DR InterPro; IPR000906; ZUS-1.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; tsp.1; 2.
 DR Pfam; PF00791; ZUS; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00408; IGC2; 1.
 DR SMART; SM00209; TSP1; 2.
 DR SMART; SM00218; ZUS; 1.
 DR PROSITE; PS0017; DEATH DOMAIN; 1.
 DR PROSITE; PS0835; IG LIKE; 1.
 DR PROSITE; PS0092; TSP1; 2.
 KW Immunoglobulin domain
 SQ SEQUENCE 945 AA; 103725 MW; 43D33B4524E0CBF2 CRC64;

Query Match 52.1%; Score 2493; DB 11; Length 945;
 Best Local Similarity 52.1%; Pred. No. 3e-221;
 Matches 497; Conservative 152; Mismatches 241; Indels 64; Gaps 18;

QY 1 MAVRCLWPAALLGIVLAAM-----LEG--SGAQOSATVANPVGPANPDLLPHFLVPEOV 53
 DB 1 MARSQVRSALLALLLLCWDPTPTSLAGVDSAGQ--VLPDSTPSAPAEQLPFLLPQDA 57
 QY 54 YIVKXKPVLLVCKAVPATQIFFKNGEWYRQVDHVIERTDSSGSEPTMEVRINVSRQOV 113
 DB 58 YIVKXKPVLLVCKAVPATQIFFKNGEWYRQVDHVIERTDSSGSEPTMEVRINVSRQOV 117
 QY 114 EKVFGLSEYWCQVAMSSSGTTSKQYIRIARLKNFEQEPLEAKEVSLEQGVLCRPP 173
 DB 118 EELFGLSEYWCQVAMSSSGTTSKQYIRIARLKNFEQEPLEAKEVSLEQGVLCRPP 177
 QY 174 EGTPPAEVEMLNEDLVPSLDPNVYITREHSIVVQARLADTANTYCVAKNIVARRSA 233
 DB 178 EGVFAEVEMLNEDLVPSLDPNVYITREHSIVVQARLADTANTYCVAKNIVARRSA 237
 QY 234 SAAVIVYVNGGWSMTWTSVCSACGRGWKRSCTNPAPLNGAFPCGQNVHRTVSS 293
 DB 238 AATVIVYVNGGWSMTWTSVCSACGRGWKRSCTNPAPLNGAFPCGQNVHRTVSS 296
 QY 294 LLVSVDGWSWPKWSACGLDCTHWSRSCSDPAPRNGGECOGTDLDTNCTSDLCV-- 351
 DB 297 TVCPVDGAWTEWSKWSACSTCAHWSRSCSDPAPRNGGECOGTDLDTNCTSDLCV-- 356
 QY 352 -----HSASGPEPDAVYVGL--IAVAVCLVLLVLLVYCRKKEGLSDVDASS-I 400
 DB 357 QRTLNDPKSHPLETSGDVALYAGLVAVFVAVVAVLMAEGLVYVRRNCRDFDTIDSSAA 416
 QY 401 LTSGFQVSIKPSKADPNHLL--TIQDPLSTTT--TYQGSICPRQDGPSPKQFQTNHLLS 457
 DB 417 LTGFGFHVNFKTARPNNPOLLHPSAPPDLTASAGIYRGPVVALQDS--ADKIPMTNSPLD 475
 QY 458 PLOGGHTLHSS-----PTSAEEFVSRLSTQNYERS----- 490
 DB 476 PLPSLKIKNYSSTIGSGGLADGALLGVLPCTYPGDF--SRDTHFLHRSASLSGQHL 534
 QY 491 --LPRGTNMTYGTENFLGGRMLPNTGISLLIPDAIPRGKIVEIYLLHKPDVRLPL 548
 DB 535 LGLPDPDSSVSGTFCGLGRSLPDTGVSLVWNGAIPOGKFYDLYLHINKAEST-LPL 593
 QY 549 A-QCOTLLSPVSCGPPGVLLTTPVILAMDHCGEPSPDSWSLRLKQSCGSGWEQDVHL 607
 DB 594 SEGSIQVLSVSVTCGPTGLLCRPFVLTVPCHCAEVIAGDWIFOLKTAHQGHWE-EVVT 652
 QY 608 GEAPSHLYCYOLEASACVFTQGLGFALVGALSAVAARLKLILLFAPVACTSLEYNI 667
 DB 653 DEETLNTPCYQLEAKSCHILLDQLGSYVFMGESYSKSAVKRLQALFALPACTSLEYSL 712
 QY 668 RVYCLDHTDALKEVVOLEKQGLQIOEPRVLHFKDSYHNRLSLIHDVPSLSKSLV 727
 DB 713 RVYCLEDTFVALKEVLELETLGVLVEEPKPLFKDSYHNRLSLIHDVPSLSKSLV 772
 QY 728 SYQEIPIFYHIWNGTORYLHCTFTLERVSPSTSLACKLWVQVGEQGSISFNITKOT 787
 DB 773 KYQEIPIFYHVWNGSQRALHCTFTLERHSLASTFTCKVCVQVQVEGEQIIFQLHTTLA-ET 831
 QY 788 RFAELIALESEAG--VPALVGPSAFKIPFLIRQKIISLDPDPCERGDWTLQAKHLDS 845
 DB 832 PAGSLDALCSAPGNATITQIGPYAFKIPFLIRQKICSSLDAPDSRGNWMLLQAKLSMDR 891
 QY 846 HLSFFASKSPPTAMINLWEARHPNGLNSQLAAAVAGLQPDAGLFTVSEAC 899
 DB 892 YLNYFATKASPTGVILDLWEARQDDGLNSLASALEEMKSEMLVAMATDGC 945

RESULT 13

Q8IZJ1 PRELIMINARY; PRT; 934 AA.
 AC Q8IZJ1;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DR	PROSITE, P550092; TSP1; 2.	QW	SEQUENCE	945 AA; 103637 MW; 56064E35F323447 CRC64;	DB 4; Length 945;	Indels 249;	Mismatches 220;	Gaps 18;
QY	Query Match							
QY	Best Local Similarity							
QY	Matches 493; Conservative 150;							
QY	1 MAYRPGWLPALGIVIAAM---LRSGGAQ--SATVANVPVGPANPDLLPHFLVEPDDYIV 56							
QY	1 MGARSARGALLIALLCWDRPLSQAGTDSGSEVLFPSPAPAPLEPYFLQEPQDAYIV 60							
QY	57 KNPVLLVCKVAPATQIFPKCNGEWVQVDHVIERSDTGSSCEPTMEVIRINVSQQVEKV 116							
QY	61 KNPVLRCKRAFPATQIYFKCNGEWVSQNDHVTQGLDEATGLRVQIEVSQQVEEL 120							
QY	117 FGLEWYWCQVWSSSGTTKSKAYIRIARLKNFQEPFLAKVSELEQGIIVLPCRPBEGI 176							
QY	121 FGLEWYWCQVWSSAGTTKSRAYVRIARLKNFQEPFLGKEVPLDHEVILLQCRPPGV 180							
QY	177 PPAVEWLRNEDLVDSLDPNVITREHSLVVRQALADTANTYTCVAKNIIVARRSASAA 236							
QY	181 PVAEVEWLRNEDLVDSLDPNVITREHSLVVRQALADTANTYTCVAKNIIVARRSASAA 240							
QY	237 VIVVNGWSTWTSWVSCSACSGRGWQKRSCTNPAPLNGGAFCEGQVNVHRTVTSLLV 296							
QY	241 VIVVNGWSTWTSWVSCSACSGRGWQKRSCTNPAPLNGGAFCEGQVNVHRTVTSLLV 299							
QY	297 SVDGSPWKSACGLDCTHWRSCSDPAPRNGGECQGTDLTRNCTSDLCVHS--- 353							
QY	300 PVDGATWKSACSTECARHWSRECMAPPPQNGRGDCSTGLDSKNCSTGLCMQNKKT 359							
QY	354 -----ASGPDVALYVGL-IAVAVCLVLLVLLVLCVCKEGLDSDVADSS-IL 401							
QY	360 LSDNSHLLASG---DAALVAGLVVAVFVVVAILMAVGVVVRNCRDFTDITDSSAAL 417							
QY	402 TSGFQPSVSKSKADNPHLL--TIQPDLSITT--TYGSLCPRGDQSPKQFTNGHLLSP 458							
QY	418 TGGFHPVNFKTARSPNQQLLHPSVPDLTASAGIYRGPVVALQDS--TDKIPMTNSPLLP 476							
QY	459 LGGGRHTLHSSPT-----SEAEFVSRLSTQNY-----FES----- 490							
QY	477 LPSLKVKYVYSSTTSGSGPLDAGDLGLVPPGTPSPDFARDTHFLHLSASLGSQLLG 536							
QY	491 LPRGTSNMTYCTFNFLGRLMIPTNGISLIPDPAIPRGKIYIYTLHKPDVRLPLA- 549							
QY	537 LPRDGSVSGTFCGLGRSLIPGTGSLVLPNGAIPQGFYEMYLINKAEST-LPLSE 595							
QY	550 QCQTLLSGTIVSCGPPGVLLTRPVILAMDHGCEPSDWSLRLLKQSCGEGSWQDVULHGE 609							
QY	596 GTQTVLSPSVTCGPTGLLCKRPVLTTPHCAEVSARDWIPQLKTAHQGHWE-EVYTLDE 654							
QY	610 EAPSHLYVQLEASACVYFTQLGRFALVGEALSVAARLKLLEAPVACTSLSEYNIRV 669							
QY	655 ETLNTPCYQLEPRACHILLDQGTGYFTGESYRSNAVKELQALAVAPALCTSLVSLRV 714							
QY	670 YCLDHTDALKEVQLEKQLGGQLIQBRVILFKDSYHNLRSLIHDVPVSLKSKLLVSY 729							
QY	715 YCLEDTPVALKVELELERTLGLYVLRPKLPMFKDSYHNLRSLIHDLPRAHWRSLKARY 774							
QY	730 QIIPYHITWGTORYLHCTFTLERSVSTSDLACKLWVQVEGGQSFSENFITKDTTF 789							
QY	775 QIIPYHITWGSQKALHCTFTLERSVSTSDLACKLWVQVEGGQSFSENFITKDTTF 833							
QY	790 AELLALEEAG--VPALVGPFAKIPFLIRKIISSLDPPRCRGAOWRTLAQKLHLSHL 847							
QY	834 GSLDLTLCAPGSTVITQGPYAFKIPLSIRKICNSLDAPNSRGNDRWMLAQKSLMDRYL 893							
QY	848 SFFASKPSPPTAMILNLEARHFPNGNLSQLAAVAGLQGPDAFLFTVSEAE 899							
QY	894 NYFATKASPTGVLIDLWEALQDDGLNLSALAEEMKSEMLVAVATDGC 945							

```

Db      311  VWSWVCSPECEHLRIRECTAPPNRNGKFCGLSQESENCTDGLCKLOKPLHEIKPQ 370
QY      352  ----HSASGPEVALYVGLIAVAVCLVLLVLLVLYCRKEGLDSVADSSILTSGFQP 407
Db      371  RWSRRGIENASDIALYSGL-GAAVAVAVAVIIVGTYLRRSHSDYGVVDVSSALTGGFQT 429
QY      408  VSIKPSKADNPILL--TIQDLSSTTTYQGSCLCPRODQSPKPOLTNHLLSPIG----- 460
Db      430  FNFKTVROGNSLLNPNAMQFDLIVSRITYSGPIC-LQD-PLDKELMTESSELPNLSIKVK 487
QY      461  -----GGRH-----TLHSSPTSEAEFEFVSRSTQNYER 489
Db      488  VQSSFMVSLGVSEAEYHGKXHSSTFPHGNNGRFSITHPRNKT---PYIONLS----- 537
QY      490  SLPRGTSNMTYGTNFIPLGGLMIPNTGISLLIPDAPRGKIYIYILTLHKPEDVRLPLA 549
Db      538  SLPTRELTRTGVFGLGGLVMPNTGVSLIPHGAIPENSWEIYMSINQGEPS-SLQSD 596
QY      550  GCOTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDPSWSRLKXKSCGSGWEODVLHGE 609
Db      597  GSEVLLSPEVTCGPPDMVITTPALTIHPCADVSSSEHNIHLKRTQOGKWE-EVMSVED 655
QY      610  EAPSHLYYQLEASACVYFTEQGRFALVGEALSVAAAKELKLLFAPVACTSLEYNIRV 669
Db      656  ESTS--CYCLDPPFACHVLLDSFGTVALTGEPIITDCAVKQLKVAVFGCMSCNSLDYNLRV 713
QY      670  YCLHDTHDALKEVVQLEKQIQQGLIOEPRVLHFKDSYHNLRLSIHDVPSLWKSLLVSY 729
Db      714  YCVDNTPCAFQEVISDERHQQGQLLEPKLLHFKGNTFSLQSVLDIPPPFLMRKPTAC 773
QY      730  QEIPFYHIWNGTORYLHCTFTLERSVSPSTDACKLWVQVEGDGQSFISINFNITKDTRF 789
Db      774  QEVPSERVWSSNRQPLHCAFSLERYTPTTQLSCKICIRQLKGHEQLQVQISILESEBE 833
QY      790  AELLALESEAGVPALVGPFAFKIPFLIRQKIISLDDPPCRRGADWRTLAQKLHLDLHLSF 849
Db      834  TITFFAQEDSTFPAQTGPKAFKIPYSIRQIRICATFDTPNAGKDWQMLAQKNSINENLSY 893
QY      850  FASKPSPTAMILNLEARHPFNGNLSQALAAVAGLQOPDAGLFTVSEAE 898
Db      894  FATQSSPSAVILNLEARHQDQDGLDSLACALEEIGRTHTKLSNITEPQ 942

```

Search completed: October 6, 2004, 18:19:33
Job time : 89 secs

mis Page blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 6, 2004, 18:20:14 ; Search time 5496 Seconds
(without alignments)
4884.662 Million cell updates/sec

Title: US-09-970-944-2
Perfect score: 4787
Sequence: 1 MAYRPGIWPALLGIVLAAML.....AVAGLGQPDAGLFTVSEAC 899

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-O=/cgn2.1/USPTO.spool/US09970944/runat_05102004.112010.1636/app.query.fasta_1.1095
-DB=EST -CPMP=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-USER=US09970944 @CGN 1.1 4087 @runat_05102004.112010.1636 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_estcl.*
10: gb_estc2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_fod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2686	56.1	3790	11	AK031655 Mus muscu
2	2665.5	55.7	2802	29	AY406491 Homo sapi
3	2651.5	55.4	2791	29	AY406493 Mus muscu
4	2493	52.1	3866	11	AK018177 Mus muscu
5	2269	47.4	2532	29	AY411749 Mus muscu
6	2253	47.1	2532	29	AY411747 Homo sapi
7	2169.5	45.3	2802	29	AY406492 Mus muscu
8	2055.5	42.9	2775	29	AY401471 Mus muscu
9	2050.5	42.8	2775	29	AY401469 Homo sapi
10	1890	39.5	2532	29	AY411748 Pan trogl
11	1864.5	38.9	1532	11	BC033727 Homo sapi
12	1550	32.4	2507	29	AY401470 Homo sapi
13	1373.5	28.7	1034	12	BI758231 603029876
14	1341.5	28.0	818	12	BI818609 603033362
15	1274	26.6	874	13	BQ689148 AGENCOURT
16	1268	26.3	889	13	BQ691915 AGENCOURT
17	1259.5	26.3	788	9	AI951556 wv36f04.x
18	1253	26.2	788	14	CA317532 UI-M-FW0-
19	1253	26.2	843	13	BX348193 BX348193
20	1236.5	25.8	756	13	BUE12387 UI-M-EW0-
21	1197	25.0	1175	10	BF530640 602071931
22	1192.5	24.9	751	14	CF735550 UI-M-HB0-
23	1189.5	24.8	749	14	CF735417 UI-M-HB0-
24	1186.5	24.8	723	13	BU613458 UI-M-EW0-
25	1126.5	23.5	934	10	BF311804 601897316
26	1126	23.5	909	13	BX364574 BX364574
27	1122	23.4	678	14	CA749784 UI-M-FD0-
28	1112.5	22.5	859	10	BF311896 UI-M-FD0-
29	1077	22.2	675	14	CA315487 UI-M-FW0-
30	1026	21.4	1040	13	BX345407 BX345407
31	1024	21.4	662	9	AL516580 AL516580
32	1022	21.3	843	12	BG913440 602811321
33	1016.5	21.2	604	10	BE314370 601147261
34	990.5	20.7	939	13	BX345406 BX345406
35	989	20.7	1201	13	BX422753 BX422753
36	934.5	19.5	720	9	AI959236 fQ25P10.Y
37	924	19.3	540	12	BI467799 389583.MA
38	900	18.8	824	12	BI737024 603360874
39	879	18.4	548	12	BI900830 ib81904.Y
40	859.5	18.0	1147	13	BU840446 AGENCOURT
41	856.5	17.9	1111	12	BG298307 602397080
42	849.5	17.7	987	13	BX452510 BX452510
43	846.5	17.7	514	13	BX522364 BX522364
44	827	17.3	815	14	CK140529 AGENCOURT
45	814.5	17.0	608	12	BM487397 pgm2n.pk0

ALIGNMENTS

RESULT 1
AK031655
LOCUS
DEFINITION Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
enriched library, clone:6030473H24 product:unc5 homolog (C.
elegans) 3, full insert sequence.
ACCESSION AK031655
VERSION AK031655.1 GI:26327502
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
FURNED

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--394-format
sequencing pipeline with 384 multiplex capillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3790)

ADACHI, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. 3790
/organisms="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:6030473H24"

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
FURNED

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--394-format
sequencing pipeline with 384 multiplex capillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3790)

ADACHI, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. 3790
/organisms="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:6030473H24"

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
FURNED

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--394-format
sequencing pipeline with 384 multiplex capillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3790)

ADACHI, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission

Submitted (16-J

Db 607 TATCTCGGAGACATTCGAGCAGGAACCCCTTGGAAAGGAAGTGTCTCTTGGAGCAGGA 666
Qy 166 IleValLeuProCysArgProProGluGlyLeuProAlaGluValGluTrpLeuArg 185
Db 667 GTCTTACTCCAGTGTCCGACACCTGAAGGATCCAGTGGCTGAGGTAGGATGGCTAAAG 726
Qy 186 AsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSer 205
Db 727 AATGAACACATAATTGATCTGCTGAAGATCGGAACCTTTATATATCTATCGATCACAC 786
Qy 206 LeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsn 225
Db 787 CTGATCATCAAGCAAGCCGACCTCTCAGATACAGCAAAATATATACCTGTGTGCAAAAT 846
Qy 226 IleValAlaArgArgSerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrp 245
Db 847 ATTGTTGCCAAGAGAAAACACACAGCCACTGTCTATCGTATGTTAATGTGGCTGG 906
Qy 246 SerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArg 265
Db 907 TCCACCTGGACAGAGTGTCTGTGTAAACAGCCGCTGTGGCGAGGATATCAGAAAGC 966
Qy 266 SerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsn 285
Db 967 ACAAGAACCTGCACCAACCCAGCCCACTCAATGGTGGGCCCTTCTGTGAGGGGCGAGT 1026
Qy 286 ValHisAspArgThrValSerSerLeuValSerValAspGlySerTrpSerProTrp 305
Db 1027 GTG--CAGAAATAGCATGCATCTATGTATGTCCAGTGGATGTAGGTGACTTCATGG 1083
Qy 306 SerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCysSerAsp 325
Db 1084 AGCAAAATGGTCAACCTGTGGAGTGAATGCACCCACTGGCGCAGGAGGAGTGTACAGCA 1143
Qy 326 ProAlaProArgAsnGlyGlyGluGlyCysGlnGlyThrAspLeuAspThrArgAsnCys 345
Db 1144 CAGCCCCCAAGAACGGGGTAAGGACTGTGATGGCTGTGCTCCATCAACGAAGACTGC 1203
Qy 346 ThrSerAspLeuCysValHis----- 352
Db 1204 ACTGATGGGCTGTGATGATGAGGATTCATTACCCCATTTCACTGAGCAGACACCCAG 1263
Qy 353 -----SerAlaSerGlyProGluAspValAlaLeuTyrValGlyLeu 366
Db 1264 AATGAATATGGATTTCTTCTGCTCTGCTCAGATGATGTGCTCTCTACGTGGGAT 1323
Qy 367 ---IleAlaValAlaValCysLeuValLeuLeuLeuValLeuIleLeuValTyrCys 385
Db 1324 GTGATCGCTGAACAGTCTGTCTGGCGATCACTGTGTGGGCCCTGTTGTGTATCGG 1383
Qy 386 ArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSerGlyPhe 405
Db 1384 AAGAACCACCGTGACTTTGAGTCTGACATCATTTGACTCTCTCAGCATCAATGGCGCTTT 1443
Qy 406 GlnProValSerIleLysProSerLysAlaAspAsnProHisLeuThrIleGlnPro 425
Db 1444 CAGCCTGTGAACATCAAGGCTGCCAGCAAGAT-----CTCCTGCTGTCCCCCT 1494
Qy 426 AspLeu---SerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 444
Db 1495 GACCTCACCTCAGCTGCAGCCATGTACAGGGGACCTGTCTATGTCTGTGCATG---GTC 1551
Qy 445 SerProLysPheGlnLeuThrAsnGlyHisLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1611
Db 1552 TCAGACAAATCCCAATGACCAACTCTCCAAATCTGGACCCACTACCCCACTGGAATATC 1611
Qy 465 ThrLeuHisHisSerSer-----ProThrSerGluAlaGluGluPheValSer 480
Db 1612 AAAGTGTACACAGCTCAGGTCTGTCACTCTCAGGATGACCTTGGCGAGTCTCACTCC 1671
Qy 481 ArgLeuSer-----ThrGlnAsnTyrPhe----- 488
Db 1672 AAATGTCAACCCAGATGACCCAGTCTTGTGTAGAGAATGAGGCCCTTAACTGGAAGAAC 1731

Qy 489 ArgSerLeuProArgGlyThr-----SerAsnMetThrTyrGlyThrPheAsnPheLeu 506
Db 1732 CAGAGCTCCACAGACAGACTGACCCATCTGTCACAGCATTTGGTACTCTTCAACTCTCT 1791
Qy 507 GlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIle 526
Db 1792 GGGGGTCACTCATCATCTTCAATTCAAGAGTGAAGTGTGCTGATTCCTCGCTGGGGCCATT 1851
Qy 527 ProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeu 546
Db 1852 CCTCAGGGAGAGTCTATGAATGTATGTACTGTACACAGGAAGAAATATACAGGCC 1911
Qy 547 ProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyVal 566
Db 1912 CCATGGAAAGACTCTCAGACCCCTACTTACCCCTGTGTGAGCTGTGGGCCCTCTCTGGAGCT 1971
Qy 567 LeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSer 586
Db 1972 CTGCTGACCCGCTCTCATCTCTCATCTGTGCATCTGTGCAGACCCACACCCAGGAC 2031
Qy 587 TrpSerLeuArgLeuLysLysGlnSerCysGlySerTrpGluGlnAspValLeuHis 606
Db 2032 TGAAGATCCAGCTCAAAACACAGGCAGTGCAGGGACAATGGGAG--GATGTTGGTG 2088
Qy 607 LeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyr 626
Db 2089 GTTGGGGAGAGAACTTCAACCCCTCTGTATCATTTCACTGTGATGCAGAGGCTTCCCAT 2148
Qy 627 ValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAla 646
Db 2149 ATCTCAGAGAGAACTCTAGTACCTATGCCCTGTGTGGCAGTCCACCAAGCAGCT 2208
Qy 647 AlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsn 666
Db 2209 GCCAAGCGTCTAAACTGGCCATCTTTGGGCCCTCTGCTGCTCTTCTTCCCTGGAGTACAGC 2268
Qy 667 IleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGlu 686
Db 2269 ATTAGAGTCTACTGCTGGATGACACACAGAGATGCCCTGAAGGAAGTTCTACAACTGGAG 2328
Qy 687 LysGlnLeuGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyr 706
Db 2329 AGGCAATGGGAGGACAGCTCTTAGAAGAACCAAGGCTCTTCGTTTAAAGCAGCATC 2388
Qy 707 HisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeuLeu 726
Db 2389 CACAACCTGGCGCTGTCTATTATGACATCGGCCCATTCCTCTGGAAGAGCAATGCTG 2448
Qy 727 ValSerTyrGlnGluIleProPheTyrHisIleTyrAsnGlyThrGlnArgTyrLeuHis 746
Db 2449 GCTAAGTATCAGGAATTCATTTTACCACATCTGGAGTGGCTCTCAAGAAACCTCCAC 2508
Qy 747 CysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrp 766
Db 2509 TGCACCTTCACTCTGGAAGACTCAGCTTAACACAGTGGAACTGGTTTGCATACTCTGT 2568
Qy 767 ValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAsp 786
Db 2569 GTGCGCAGGTGGAAGGAGAGGGCAGATCTTCCAGCTCACTGTACTGTGTCTCAGAGNA 2628
Qy 787 ThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValProAlaLeuValGly 806
Db 2629 CCTACTGGCATCGATTTACCTTCTCTGAGCCCTGTAGTACCATCACCAGTGTCAACCGA 2688
Qy 807 ProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspPro 826
Db 2689 CCAAGTGTCTTACAGCATTCCTCTCCCTATCCGAGAGAGTATGCAGAGCGCTGTGATGCC 2748
Qy 827 ProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLysAspSerHis 846
Db 2749 CCTCAAAACAGAGGCCATGACTGGAGGATGTGGGCCCATAACTCACTGCAGAGGTAC 2808

Qy	764	LysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsn	PheAsnIle 783
Db	2392	AAACTCTGTGTGGCGCAGGTGGGAAGAGGGCAGATCTTCCAGCTCAACTGCACCGTG	2451
Qy	784	ThrLysAspThrArgPheAlaGluLeuAlaLeuGluSerGluAlaGlyValProAla	803
Db	2452	TCAGAGAACTACTGGCATCGATTTGCCGCTGCTGGATCTCTGGAAACACCATCACCG	2511
Qy	804	LeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleSerSer	823
Db	2512	GTACCGGGGCCAGTGTCTTCAGCATCCCTCTCCCTATCCGCAGAGCTCTGTAGCAGC	2571
Qy	824	LeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeu	843
Db	2572	CTGGATCCCCCAGACAGAGGCGCATGACTGGAGGATGCTGGCCATAAGCTGAACCTG	2631
Qy	844	AspSerHisLeuSerPheAlaSerLysProSerProThrAlaMetIleLeuAsnLeu	863
Db	2632	GACAGGTACTTGAAATCTTTGGCCAAATCCAGGCCCACTGGCGTAAATCTCTGGATCTT	2691
Qy	864	TrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGly	883
Db	2692	TGGGAAGCAGAACTTCCAGATGGAACTGAGCATGCTGGCAGCTGCTCTGGAGAA	2751
Qy	884	LeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGlu	898
Db	2752	ATGGGAAGCATGAAACGGTGGTGCTCTTAGCAGCAGAGGGCGAG	2796
RESULT 3			
AY406493			
LOCUS	AY406493	2791 bp	DNA linear GSS 15-DEC-2003
DEFINITION	Mus musculus UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.		
ACCESSION	AY406493		
VERSION	AY406493.1	GI:39762467	
KEYWORDS	GSS.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferrieria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios		
JOURNAL	Science	302 (5652), 1960-1963 (2003)	
PUBMED	14671302		
REFERENCE	2 (bases 1 to 2791)		
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferrieria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment		
FEATURES	Location/Qualifiers		
source	1..2791		
gene	/organism="Mus musculus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:10090"		
	<1..>2791		
	/gene="UNC5C"		
ORIGIN	/locus_tag="HCM2575"		
Alignment Scores:			
Pred. No.:	3,76e-225	Length:	2791
Score:	2651.50	Matches:	511
Percent Similarity:	72.49%	Conservative:	153

Best Local Similarity:	55.79%	Mismatches:	214
Query Match:	55.39%	Indels:	38
DB:	29	Gaps:	14
US-09-970-944-2 (1-899) x AY406493 (1-2791)			
QY	9	ProAlaLeuLeuGlylleValLeuAlaAtrpLeuArgGlySerGlyAlaGlnGlnSer	28
DB	76	CTGCGCTG-----GCCCTGCTAAGCGAGTGGCCACCGGCTCCCGCTCAAGATGAT	129
QY	29	Ala-----ThrValAlaAsnProValProGlyAlaAsnProAspLeuLeuProHis	45
DB	130	GAATTTTTCAGAACTCCAGAAACCTTCCATCTGACCCACTGAGCAATGCCACAC	189
QY	46	PheLeuValGluProGluAspValTyrIleValIleAsnLysProValLeuLeuValCys	65
DB	190	TTCTCTATTGAGCCCGAGGAAGCTTACATTGTGAAGAACAAAGCTGTGAACCTGTATTGT	249
QY	66	LysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArgGlnVal	85
DB	250	AAAGCCAGCCCTGCCACCCAGATCTACTTCAAGTGCACACGAGTGGGTTCATCAGAAG	309
QY	86	AspHisValIleGluArgSerThrAspGlySerSerGlyGluProThrMetGluValArg	105
DB	310	GACACGTAGTACGAGAGAGTAGATGAACCTCTGCTTAATTGTGAGAGAGTGAGC	369
QY	106	IleAsnValSerArgGlnGlnValGluIleValPheGlyLeuGluGluTyrTrpCysGln	125
DB	370	ATTGAGATTTCACGACAGAGGTGGAGAACTGTTGGGCTGAAGATTACTGGTGCCAG	429
QY	126	CysValAlaTrpSerSerGlyThrThrySerGlnLysAlaTyrIleArgIleAla	145
DB	430	TGTGTGCTGGAGCTCAGCAGGACCTACGAGAGTCCGAGAGGATACGTCGCGCAITGCG	489
QY	146	-----ArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGlu	163
DB	490	TTTCCAGATCTGCGAAGACATTCGACAGAGAACCTTGGGAAGAGTGTCTTGAG	549
QY	164	GlnGlyIleValLeuProCysArgProProGlu-GlyIleProProAlaGluValGluTr	183
DB	550	CAGGAAGTCTACTCCAGTGTGGCCACCTGAAGGGA-----TCCGTAGAATG	597
QY	183	pLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgG	203
DB	598	GCTAAAGAATGAACACATAATTGATCTGCTGAAGATCGGAACCTTTTATATATCTGA	657
QY	203	uHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAl	223
DB	658	TCACAACCTGATCATCAGCAGCCGACTCTCAGATACAGCAATATATACCTGTGTGC	717
QY	223	aLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyrValAsnG	243
DB	718	CAAAAAATATTGCGCAAGAAAAAGCACCACACCACTGCTCATCTGTATGTATATGG	777
QY	243	YGLYTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpG	263
DB	778	TGGTGTGTCCACCTGGACAGAGTGTCTGTGTAAACAGCCGCTGTGGCGGAGATATCA	837
QY	263	nLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluG	283
DB	838	GAACGACACAGAAGCTGCACCAACCCAGCCCACTCAATGGTGGGCTTCTGTGAGGG	897
QY	283	YGLAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrpSe	303
DB	898	GCAGAGTGTG---CAGAAAAATAGCATGCACTGATATGTCCAGTGGATGTAGTGGAC	954
QY	303	rProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCy	323
DB	955	TTCATGGAGCAANTGTCACTGTGGAGTGAATGCCACCACTGCGCGCAGAGGAGTG	1014
QY	323	sSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrAr	343
DB	1015	TACAGCACCGCCCCCAAGAACGGGGTAAAGGACTGTGTATGGCTGTCTCCAATCCAA	1074

QY	343	gAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTy	363
DB	1075	GAATGCACTGATGGCTGTGATCAGAGCTGCTCTGACTCAGATGATGGCTCTCTA	1134
QY	363	rValGlyLeu-----IleAlaValValCysLeuValLeuLeuLeuValLeuIleLe	382
DB	1135	CGTGGGATGTGATCGCTGTAAACAGTCTGTGGCGATCATCTGTGTGGCCCTGT	1194
QY	382	uValTyrCysArgLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuTh	402
DB	1195	TGTGTATCGGAAGAACCCCGTACTTTGAGTCTGACATCATTCATCTCTCAGACTCAA	1254
QY	402	rSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeuTh	422
DB	1255	TGGCGGCTTTTCAACCTGTGAACATCAAGGCTGCCACAGAT-----CTCCTGGC	1305
QY	422	rIleGlnProAspLeu---SerThrThrThrThrGlnGlySerLeuCysProArgG	441
DB	1306	TGTCCTCCCTGACCTCACCTCAGCTCAGCCATGTACAGGGGACCTGTCTATGCTGCA	1365
QY	441	nAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyG	461
DB	1366	TGAT---GTCTCAGACAAAATCCCAATGACCAACTCTCCNAITCTGACCCACTACCAA	1422
QY	461	YGLYArgHisThrLeuHisSerSer-----ProThrSerGluAlaGluG	477
DB	1423	CTTGAATAAATAAGTGTACACAGCTCAGGTGTGTCTCTCTCAGGATGACCTTGCCGA	1482
QY	477	uPheValSerArgLeuSer-----ThrGlnAsnTyrPhe-----	488
DB	1483	GTCTCATCAAACTGTACCCAGATGACCCAGTCTCTGTAGAGAAATGAGGCCCTTAA	1542
QY	489	-----ArgSerLeuProArgGlyThr-----SerAsnMetThrTyrGlyThrPh	503
DB	1543	CCTGAAGAACCCAGAGCCCTCGCAAGACAGACTGACCCATCTCTGACACATTTGTACTT	1602
QY	503	eAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIleProPr	523
DB	1603	CAACTCTCTGGGGTCACTCATCTCTTCTTATCAGGAGTAAGTTGCTGATTCGCCG	1662
QY	523	oAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAs	543
DB	1663	TGGGGCCATTCTCTCAGGGAGAGTCTATGAATGTATGTACTGTACACAGGAAAGAAA	1722
QY	543	pValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyPr	563
DB	1723	TATGAGCCCTCCATGGAAGACTCTCAGACCCCTACTTACCCCTGTGTGAGCTGTGGCC	1782
QY	563	oProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSe	583
DB	1783	TCTTGGAGCTCTGTGTACCCGCTGTCTCTCTCTGATCTGATCTGATGTCAGACCCAG	1842
QY	583	rProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGlySerTrpGluGlnAs	603
DB	1843	CACCGAGACTTGAAGATCCAGCTCAAAACACAGGAGTGCAGGACAAATGGGAG---GA	1899
QY	603	pValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSe	623
DB	1900	TGTTGTGGTGTGGGAGGAGAACTTCACACCCCTGTGTATTCAGCTGGATGTCAGA	1959
QY	623	rAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSe	643
DB	1960	GGCTTGCCATCTCTCAGAGAACTTCAGTACCTATGCTGCTGTGGGACAGTCCACAC	2019
QY	643	rValAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLe	663
DB	2020	CAAAAGCAGCTGCCAAGCGTCTTAAACTGGCCATCTTTGGGCCCCCTCTCTCTTCCCT	2079
QY	663	uGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValVa	683
DB	2080	GGAGTACAGCATTAGCTCTACTGCTGATGACACACAGGATGCGCTGAGGAAGTTCT	2139


```
Qy 693 lGlnLeuGlnLysGlnLeuGlyGlyGlnLeuLeuGlnGluProArgValLeuHisPheLy 703
Db 2140 ACAACTGAGAGGAAATGGGAGGACAGCTCTAGAGAACCCCAAGGCTCTTCATTAA 2199
Qy 703 sAspSerTyHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSe 723
Db 2200 AGGAGAGATCCAACTGGCCCTGTCTATTTCATGACATCGCCCATTCCTCTCGAAGAG 2259
Qy 723 rLysLeuLeuValSerTyGlnGlnLleProPheTyHisIleTrpAsnGlyThrGlnAr 743
Db 2260 CAAATTGCTGCTAAGTATCAGGAATTCATTTTACACATCTGGAGTGGCTCTCAAG 2319
Qy 743 gTyLeuHisCysThrPheThrLeuLeuArgValSerProSerThrSerAspLeuAlaCy 763
Db 2320 AAACCTCCACTGCACCTTCACTCTGGAAGAGCTCAGCTTAACACAGCTGGAACCTGTTG 2379
Qy 763 sLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnII 783
Db 2380 CAAACTCTGTGGCGGAGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2439
Qy 783 eThrLysAspThrArgPheAlaGlnLeuLeuAlaLeuGluSerGluAlaGlyValProAl 803
Db 2440 GTCAGAGAACCTTACTGGCATCGACTTACCTCTCTGGACCTCTGCTAGTACCATCACAC 2499
Qy 803 aLeuValGlyProSerAlaPheLysIleProPheLeuLeuArgGlnLysIleIleSerSe 823
Db 2500 TGTCAACGGACCAAGTCTTTCAGCATCTCTCCCTATCCGCGAGAAGCTATGCAGCAG 2559
Qy 823 rLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLe 843
Db 2560 CTTGGATGCCCTCAACAAGAGGCCATGATCTGGAGGATCTGGCCCAATAAATCAACCT 2619
Qy 843 uAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLe 863
Db 2620 GGACAGTACTTGAATTACTTTCACCAATCGAGCCCAATCGGCGTGAATCTTGATCT 2679
Qy 863 uTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGl 883
Db 2680 TTGGGAAGCACAGAACTTCCAGATGGAACTGAGCATGCTGGCAGCGTCTCTGGAGA 2739
Qy 883 yLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGlu 898
Db 2740 AATGGGAAGACATGAGACAGTGTCTTGGCAGCAGAGGACAG 2785

RESULT 4
AK018177
LOCUS
DEFINITION
Mus musculus adult male medulla oblongata cDNA, RIKEN full-length
UNC5H2 homolog [Rattus norvegicus], full insert sequence.
ACCESSION
AK018177
VERSION
AK018177.1 GI:12857775
KEYWORDS
HTC; CAP trapper
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
REFERENCE
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11042159
REFERENCE
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
```

Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matakiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

REFERENCES

4

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

REFERENCES

5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3866)

REFERENCES

6

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirakawa, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, K., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission

REFERENCES

7

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTTAATAATTAATCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

FEATURES

source

1..3866
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:6330415E02"
/db_xref="MGI:1910526"
/db_xref="taxon:10090"
/clone="6330415E02"
/sex="male"
/tissue type="medulla oblongata"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"

CDS	417..3254	/note="unnamed protein product; TRANSMEMBRANE RECEPTOR UNCSH2 homolog [Rattus norvegicus] (SPTR 008722, evidence: PASTY, 96.5%id, 100%length, match=2835) putative"
	/protein_start=1	/codon_start=1
	/db_xref="GI:12857776"	/db_xref="BAB31108.1"
	/translation="NRASGVSALLALLLLCWDPTPSLAGVDSAGQVLPDSYPSAPA EQLPYELLBPQDAIVNKPVELHCRAPATQIYFKCNGEWISQNDHVTQESLDEATG LRVREVOIEVSQVEELFGLBEDYWCQCVAMSSGTTKRRAYIRIAYLRKFDQEP L AKEPDLHVLVLCQBPPEGVPAVEMLKNEVIDIDPAQDTNELLTDHNLIRQARLS DRYNYFVAKNIVAKRSTAAVIVYNGSGSWAWEWSCNRCGRGOKRKTCTNP APNGACFEGAFQACTTTCVVDGAWTEWSKMSACSTECCHWRSRCMAFPQNG GRDCSTLNDKNDTGLCVLNQRTLNDPKSHPLETSGDVALYAGLVAVFVVAVLM AEGVIYRNRCDFDFTDSSNAALTGGFHPNFKTARENPNQLLHPSAPDLTASAG IYRGFVALQDSADKIPMTNSPELLDPLSLKI KVNYSSTIGSGSLADGADLIGLPP GTYPGDFSDTHPLHRSASLQHLGLPRDPSVSGTFCGLGRLSLPGTGVSL L VNGAIPQKFDVLYLHINKASTLPLSGSQTVLSPVTCGPTGLLCPVVLTVPH CAEVIAGDWIFOLKQAHQGHWEVVDLEETLNTFCYQLENKASHILLDQLGYSVF MGESYSRAVKRLQALFAPALCTSEYLSRVLEDTPVALKEVLELERTGLGGLVE EPKPLFKDSYHNLRLSLHDIPAHWRSLKAKYQEI PFYHWNQSQRALHCTFLER HSLASTFECKYKCVROVEGGOIFQHLHTLAETPAGSLDALCSAPGNALTITOLGPYAF KIPLSRQIKCSLDPDSDRGNDWRLLAOKLSMDRYLNYFATKASPTGVILDWEARQ QDDGDLNLSALAEEMGKSEMLVAMATDGC"	
ORIGIN		
Alignment Scores:		
Seq. No.:	9.95e-211	Length: 3866
Score:	2493.00	Matches: 497
Percent Similarity:	68.03%	Conservative: 152
Best Local Similarity:	52.10%	Mismatches: 241
Query Match:	52.08%	Indels: 64
DB:	11	Gaps: 18
US-09-970-944-2 (1-899) x AK018177 (1-3866)		
Qy	1	MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaIaTrp--- 19
Db	417	ATGAGGCGCGAGCGGGTGGCGAGCGCGTGTCTGTCGGCGTGTCTTGTGGAT 476
Qy	20	-----LeuArgGly-----SerGlyAlaGlnGlnSerAlaThrValAlaAsn 33
Db	477	CCGACACCGAGCGCTAGCAGCGGTTGACTCTGTCTGGCCAG-----GTGCTCCAGAC 527
Qy	34	ProValProGlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspVal 53
Db	528	TCCTACCATCAGCCCTCGGAGCAGTGCCTGACTTCTTATTGGAGCCACAGAGCC 587
Qy	54	TyrIleValLysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIle 73
Db	588	TACATCGTAAGAACAAAGCCAGTGGAACTGCACCTGCAGAGCTTCCCTGCCACAGATC 647
Qy	74	PhePheLysCysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThr 93
Db	648	TACTTCAAGTGTATGGCGAGTGGTTCAGCCAGATGACACAGTCACACAGAGAGCTG 707
Qy	94	AspGlySerSerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnVal 113
Db	708	GATGAGGCCACAGGCTTCCGGGTGGAGAGGTGCAGATCGAGGTGTCTACCGGCAAGTG 767
Qy	114	GluLysValPheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSerSerGly 133
Db	768	GAGGAACCTCTTCGGGCTCGAGACTACTGTGTGCCAGTGGCTGGCTGGAGCTTTCGGA 827
Qy	134	ThrThrLysSerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGln 153
Db	828	ACTACCAAGAGTCGCGAGCCTTACATCCGATGCTGCTACTTGGCAAGAACTTTGACCAG 887
Qy	154	GluProLeuAlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProPro 173
Db	888	GAGCCTCTGGCCACAGAGGTACCTTGGATCATGAGGTCTCTTCTGCTGCGCGCCACCG 947

Db 2019 CTGGCCCTACCTCGGAGCCAGCAGTGTCTGAGCGGACCTTTGGTTGCTCGGAGGA 2078
 Qy ArgLeuMetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArg 528
 Db 2079 AGGTGACCTCCCGGACAGGGTGTGAGCTGTGGTACCAATGAGGACCATTTCCCGAG 2138
 Qy GlyIleIleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProLeu 548
 Db 2139 GGCAAGTCTATGACCTATCATATCAATCAACAGGCGGAGACCC--CTCCACTT 2195
 Qy Ala---GlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyValLeu 567
 Db 2196 TCAGAAGTTCCTCCAGACAGTATGAGCCCTCGGTGACCTGTGGGCCACAGGCTACTC 2255
 Qy LeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrp 587
 Db 2256 CTGTGGCCCTGTGTCTCTACCGTGTCCCTGCTGTGAGTCACTGCTGGAGACTGG 2315
 Qy SerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeu 607
 Db 2316 ATCTTTAGCTCAAGACCCAGGCCATCAGGGCCACTGGAG---GAGGTGTGACTTGG 2372
 Qy GlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrVal 627
 Db 2373 GATGAGGAGACCTTCACACACCCCTGCTACTGCCAGCTGAGGCTAAAGTCTGCGCATC 2432
 Qy PheThrGluGlnLeuGlyIleArgPheAlaLeuValGlyGluAlaLeuSerValAlaAla 647
 Db 2433 CTGCTGACACGCTGGTGTCTTCTAGTATTCATGGCCGAGTCTTCTGCTGCGAGTC 2492
 Qy LysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIle 667
 Db 2493 AAGCGGTCTCAGCTGGCCATCTTCGCCCCAGCCCTGTCACCTCCCTGGAGTATAGCTC 2552
 Qy ArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLys 687
 Db 2553 AGGTCTACTGTCTGGAGGACACCTGTAGCCTGAAGGAGTCTTGGAGCTGGAGAG 2612
 Qy GlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHis 707
 Db 2613 ACTCTGGTGGCTACTTGTGGAGGAGCCCAAGCCCTTGTCTTTAAGGACGTTACCAC 2672
 Qy AsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeuLeuVal 727
 Db 2673 AACCTACGCTCTCCCTCCATGATACATCCCATGCCCCTGAGGAGGACAACTACTGGCC 2732
 Qy SerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCys 747
 Db 2733 AGTACAGAGATCCCTCTTACCACTGTGGAATGGCAGCAGAGCCCTGCACTGC 2792
 Qy ThrPheThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrpVal 767
 Db 2793 ACTTTCACCTCGAGAGGATACCTCGGCTCCAGGAGTTCACCTGTAAAGTCTCGGTG 2852
 Qy TrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThr 787
 Db 2853 CGGAGTCCGAGGGAAGCCAGATTTTCCAGCTGCACACACGTTGGCC---GAGACG 2909
 Qy ArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly-----ValProAlaLeuVal 805
 Db 2910 CTGCTGGCTCCCTGGATGCTCTGCTGTGCGCGGCAATGCCATACCACCCAGCTG 2969
 Qy GlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAsp 825
 Db 2970 GACCTTATGCTTCAAGTATACCTCTGCTCCATCCGCAAAAGATCTGCAGAGCTGGAC 3029
 Qy ProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSer 845
 Db 3030 GCCCCGACTCCCGGGCAACGACTGGAGCTGTGGCGCAGAGCTGTCCATGACCGG 3089
 Qy HisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGlu 865
 Db 865

Db 3090 TACCTAACTACTTTCGCCACCAAGCTAGTCCACAGGTGTCTTATAGACTCTGGGAA 3149
 Qy 866 AlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGly 885
 Db 3150 GCTCGGCAACAGGATGACGGGACTCAACAGCTGCGGAGTGTCTGGAGGATGGC 3209
 Qy 886 GlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
 Db 3210 AAGAGTGTGAGTGTGGTAGCCATGCCACAGATGCGGATTGC 3251

RESULT 5
 AY411749
 LOCUS
 DEFINITION Mus musculus HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY411749
 VERSION AY411749.1 GI:39767717
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE 1 (bases 1 to 2532)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene tricos
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 2532)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 Location/Qualifiers
 source 1..2532
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 gene <1..>2532
 /locus_tag="HCM4327"

Alignment Scores:
 Pred. No.: 4,94e-191 Length: 2532
 Score: 2269.00 Matches: 446
 Percent Similarity: 69.23% Conservative: 139
 Best Local Similarity: 52.70% Mismatches: 206
 Query Match: 47.40% Indels: 54
 DB: 29 Gaps: 15

US-09-970-944-2 (1-899) x AY411749 (1-2532)

Qy 103 GluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluTyr 122
 Db 13 GAGGTGCGATGCGAGTGTACGGCAGCAAGTGGAGAACTCTCGGGCTCGAGACTAC 72
 Qy 123 TrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAlaTyrIle 142
 Db 73 TGGTCCAGTGGCTGGCTTTCGGGAACTTCCAGAGTCCGCGACCTACATC 132
 Qy 143 ArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeu 162
 Db 133 CGCATTTGCTACTTTCGCGCAAGAACTTTTCCAGGAGCTCTGGCCAGGAGTACCTTG 192
 Qy 163 GluGlnGlyIleValLeuProCysArgProGluGlyIleProProAlaGluValGlu 182

1264	CTGCGCAGTGGCAGCGCTTGGTTCCGACGACCTCTCTGGGCGCTTACCTCTGGGAGCCCGGAGCAGC	1323
498	MetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIle	517
1324	ATGTCTCAGGGGACCTTTGGTTGGCTGGAGAGAGGCTGAGCCTCCCGGCACAGGGGCTC	1383
518	SerLeuLeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThr	537
1384	AGCGTGTGGTACCAAAATGGAGCCATTCCCGAGGCAAGTTCTATGACCTGTATCTACAT	1443
538	LeuHisLysProGluAspValArgLeuProIleAla---GlyCysGlnThrLeuLeuSer	556
1444	ATCAACAGAGCCGAAAGCACC---CTCCCACTTTTCAAGAGGTCCAGACAGTATTGAGC	1500
557	ProIleValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMet	576
1501	CCCTCGGTGACCTGTGGGCCACAGGCTACTCTCTGTCGCGCCCTGTCCTCACCGTG	1560
577	AspHisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCys	596
1561	CCCCACTGTGCTGAAGTCATCGCTGGAGACTGGATCTTTCAGCTCAAGACCCAGGCCCAT	1620
597	GluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyr	616
1621	CAGGGCCACTGGGAG---GAGGTGGTACCTTGGATGAGAGACCCCTCAACACCCCTGC	1677
617	TyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAla	636
1678	TACTGGCAGCTGAGGCTTAAGTCTCGCCACATCTCTGCTGGACCGCTGGTACCTACGTA	1737
637	LeuValGlyGluAlaLeuSerValAlaAlaLysArgLeuLeuLeuLeuPheAla	656
1738	TTTCATGGCGAGTCTTACTCTCGCTCGCAGTCAAGCGGCTCCAGCTGGCCATCTTCGCC	1797
657	ProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHis	676
1798	CCAGCCCTCTGCACCTCCCTGGAGTATAGCTCAGGGTCTACTGTCTGGAGGACACCT	1857
677	AspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGlnLeuIleGlnGlu	696
1858	GTAGCACTGAAGAGGTCCTGGAGCTGGAGAGGACTCTGGTGGTCTTCTGGTGGAGGAG	1917
697	ProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspVal	716
1918	CCCAAGCCTTGTCTTAAAGACAGTATACCACACTACGCTCTCCCTCCATGACATC	1977
717	ProSerSerLeuTrpLysSerLysLeuValSerTyrGlnGluIleProPheTyrHis	736
1978	CCCCATCCCCCTGGAGGAGCAAACTACTGGCCCAAGTACACGAGGATTCCCTCTTACCAC	2037
737	IleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerPro	756
2038	GTCTGGAAATGGACCCAGAGCCCTGCACTGCACTTTTCCCTTGAGAGGATAGACCTG	2097
757	SerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSer	776
2098	GCCTCCAGGAGTTCACCTCTAAGTCTGCGTCGCGCAGGTCGAGGGGCAAGCCAGATT	2157
777	PheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGlu	796
2158	TTCCAGCTGCACACAAAGTTGGCC---GAGACGCCCTGCTGGCTCCCTGGATGCTCTGC	2214
797	SerGluAlaGly-----ValProAlaLeuValGlyProSerAlaPheLysIleProPhe	814
2215	TCTGCCCGGGCAATGCGATCACCCAGCTGGACCCCTATGCTTCAAGTATACCCCTG	2274
815	LeuIleArgGlnLysIleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrp	834
2275	TCCATCCGCCAAAAGATCTCGACGAGCTCGACGCGCCCACTCCCGGGGCAACGACTGG	2334
835	ArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysPro	854
2335	AGGCTGTGGCGCGAAGAGTGTCCATCGACCGGTACTTAACTACTTCGCGCACCAAGCT	2394
193	GATCATGAGGTCTCTTCAGTGGCGCCGCCACCGGAGGAGTGCCTGTGGCTGAGGTGAA	252
183	TrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArg	202
253	TGGCTCAAGATGAAGATGTCATTGACCCGCTCAGGACACTAACTTCTGCTCACCAT	312
203	GluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysVal	222
313	GACCACAACTCATCATCCGCGAGGCGCCCTCTCAGACACGCGCCAACTACACCTGTGTG	372
223	AlaIysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyrValAsn	242
373	GCCAAAGATATCGTGCCCAAGCCCGGAGCACCCAGGCCACAGTCACTGCTATGTGAAT	432
243	GlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrp	262
433	GGAGGCTGGTCCAGCTGGGCAGAGTGGTCACCCCTGTTCCTCAATCGTGTGGCGAGGCTGG	492
263	GlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGlu	282
493	CAGAGCGTACTCGGACCTGCACCAATCAGCCGCCACTCAATGGAGGCGCTTCTGTGAG	552
283	GlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrp	302
553	GGACAG---GCCTTCCAGAGACAGCTTGCCACCCGCTGTCGCCAGTGGAGCGGTGG	609
303	SerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGlu	322
610	ACCAGTGGAGCAAGTGGTCTGCTCGACGACAGAGTGTGGCAGCTGGCGCAGCGCAG	669
323	CysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThr	342
670	TGCATGCGACCGCCACCCAGAACGAGGCGGTGACTGTCAGCGGCGGACGCTACTTCACTCC	729
343	ArgAsnCysThrSerAspLeuCysVal-----	351
730	AAGAACTGCATGTATGGGCTGTGCGTCTGCTGAATCAGAGAACTTAAACGACCTTAAAGC	789
352	HisSerAlaSerGlyProGluAspValAlaLeuTyrValGlyLeu---IleAlaValAla	370
790	CACCCCTGGAGACATCGGAGATGTGGCACTGTACGAGGCTTGTGGTGGCGGCTTT	849
371	ValCysLeuValLeuLeuLeuValLeuIleLeuValTyrCysArgLysLysGluGly	390
850	GTGGTGGTAGGGTCTCATGCGCGTGGGAGTGATCGTATACCGGAGAACTGCCGGGAC	909
391	LeuAspSerAspValAlaAspSerSer---IleLeuThrSerGlyPheGlnProValSer	409
910	TTTCAGACGACATCACCGACTCTCTGCGGCCCTCTGACCCGCTCCACCTGTCTCAAC	969
410	IleLysProSerLysAlaAspAsnProHisLeuLeu-----ThrIleGlnProAspLeu	427
970	TTCAAGACTGAAGGCCCAACACCCGAGCTCTCTGACCCGCTCCGCGCCCTCCAGACCTA	1029
428	SerThrThrThr---ThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerPro	446
1030	ACGGCCAGTGTGGCATCTACCGGGGCGCTGTGTATGCCCTTGCAGGACTCC---GCCGAC	1086
447	LysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyArgHisThrLeu	466
1087	AAGATCCCATGACTAATTCGCCCTGTGTGATCCCT	


```

488  ---PheArgSer-----LeuProArgGlyThrSer 496
1261 CACCTGGCAGCCAGCCTCGGTCCAGCAGCTCTTGGGGCTTGGCCGAGACCCAGGG 1320
497 AsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGly 516
1321 AGCAGCGCTCAGCGGCACCTTGTGCTGCTGGGTGGGAGGCTCAGCATCCCGGCACAGGG 1380
517 IleSerLeuLeuIleProProAsnAlaIleProArgGlyLysIleTyrGluIleTyrLeu 536
1381 GTACGCTTGTGTGGCCCATCGAGCCATTCCCGAGGCAAGTCTTACGAGATGTATCTA 1440
537 ThrLeuHisLysProGluAspValArgLeuProIleuAla---GlyCysGlnThrLeuLeu 555
1441 CTCATCAACAGCAGCAAGATACC---CTCCCGCTTTCAGAAGGGACCCAGACAGTATTG 1497
556 SerProIleValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAla 575
1498 AGCCCTCGGTGACCTGTGACCCACAGGCTCTCTGTGTGGCCCGCTCATCTTCACC 1557
576 MetAspHisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSer 595
1558 ATGCCCTCAGTGTGCGAAGTCAGTGCCTGCTGCTGCTTTCAGCTCAAGACCCAGGCC 1617
596 CysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeu 615
1618 CACCAGGCGCCTGAGGAG---GAGTGTGTGACCTGGATGAGGAGACCTGGAACACACCC 1674
616 TyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPhe 635
1675 TGCTACTGCCAGCTGGAGCCAGGCGCTGTGCATCTCTGTGGACACAGCTGGGACCTAC 1734
636 AlaLeuValGlyGluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuPhe 655
1735 GTGTTTCAGGCGAGTCTATTCCGCTCAGCAGTCAAGCGCTCCAGCTGCGCGTCTTC 1794
656 AlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThr 675
1795 GCCCGCCGCTCTGCACCTCGCTGGAGTACAGCTCCGGGTCTACTGCTGGAGGACAG 1854
676 HisAspAlaLeuLysGluValGlnLeuGlyLysGlnLeuGlyGlyGlnLeuIleGln 695
1855 CCTGTAGCAGTGAAGAGTGTGAGCTGGAGCGGACTCTGGCGGATCTTGTGGAG 1914
696 GluProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAsp 715
1915 GAGCGGAACCGCTAATGTTCAAGGACAGATTACCACACCTGCGCTCTGCTCCATGAC 1974
716 ValProSerSerLeuTrpLysSerLysLeuValSerTyrGlnGluIleProPheTyr 735
1975 CTCGCCCATGCGCATTTGAGGAGCAAGTGTGCGCAATACCAAGAGATCCCTTCTAT 2034
736 HisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSer 755
2035 CACATTTGGAGTGGCAGCAGAGGCTCTCAGCTTTCACCTCGAGAGGACACAGC 2094
756 ProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGln 775
2095 TTGGCTCCACAGAGCTCACCTGCAAGTCTGCTGGCGCAAGTGGAGGGGAGGCGCAG 2154
776 SerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeu 795
2155 ATATTCCAGCTGCATACCATCTGGCA---GAGACACCTGCTGGCTCCCTGACACTCTC 2211
796 GluSerGluAlaGly-----ValProAlaLeuValGlyProSerAlaPheLysIlePro 813
2212 TGCTTGCCCTGGGAGACTGTCAACCCAGCTGGGACCTTATGCTTCAAGATCCCA 2271
814 PheLeuIleArgGlnLysIleLeuSerSerLeuAspProProCysArgArgGlyAlaAsp 833
2272 CTGTCCATCGCCAGAGATATGCAACAGCTAGATGCCCCCACTCACTCGGGCAATGAC 2331

```

834 TrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLys 853

2332 TGGCGATGTTAGCAGAGAGCTCTATGAGCCGGTACCTGAATTACTTGGCACCACAA 2391

854 ProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsn 873

2392 GCGAGCCCGGCTGATCTCTGGAGCTCTGCGAGAGCTCTGCGAGGACGATGGGAC 2451

874 LeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThr 893

2452 CTCAACAGCTGCGAGTGCCTTGGAGGAGATGGCAAGAGTGCATGCTGGTGGCTGTG 2511

894 ValSerGluAlaGluCys 899

2512 GCCACCGCGGACTGC 2529

RESULT 7

AY406492 2802 bp DNA linear GSS 15-DEC-2003

LOCUS Pan troglodytes UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

DEFINITION

ACCESSION AY406492.1 GI:39762466

VERSION AY406492.1

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM

REFERENCE 1. (bases 1 to 2802)

AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2. (bases 1 to 2802)

AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES

source 1..2802

Location/Qualifiers

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/gene="UNC5C"

/locus_tag="HCM2575"

gene

ORIGIN

Alignment Scores:

Pred. No.: 4,656-182 Length: 2802

Score: 2169.50 Matches: 438

Percent Similarity: 62.95% Conservative: 138

Best Local Similarity: 47.87% Mismatches: 306

Query Match: 45.32% Indels: 33

DB: 29 Gaps: 13

US-09-970-944-2 (1-899) x AY406492 (1-2802)

9 ProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSer 28

76 CTGCGCTG-----GCCCTGCTCAGCGCCAGCTGGCTCGCGCCCAAGATGAT 129

29 -----AlaThrValAlaAsnProValProGlyAlaAsnProAspLeuLeuProHis 45

Db 130 GACITTTTTCATGAACCTCCAGAAACTTTTCTTCGGATCCACCTGAGCCTCTGCCACAT 189
Qy 46 PheLeuValGluProGluAspValTyrIleValLysAsnLysProValLeuLeuValCys 65
Db 190 TTTCTATTGAACTGAGGAGCTTATATTGTGAAGATAAGCCTGTGAACCTGTACTGT 249
Qy 66 LysAlaValPro-----AlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArg 83
Db 250 AAAGCAAGCCCTGNN 309
Qy 84 GlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluProThrMetGlu 103
Db 310 NNN 369
Qy 104 ValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrp 123
Db 370 NNN 429
Qy 124 CysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAlaTyrIleArg 143
Db 430 NNN 489
Qy 144 IleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGlu 163
Db 490 NNNNNNNNNNNCGAAAGACATTTGAGCAGAACCCCTAGGAAAGGAAGTCTCTTTGGAA 549
Qy 164 GlnGlyIleValLeuProCysArgProGluGlyIleProProAlaGluValGluTrp 183
Db 550 CAGGAAGTCTTACTCCAGTCTCGGCCACCTGAAGGATCCAGTGGCTGAGNNNNNNNN 609
Qy 184 LeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGlu 203
Db 610 NNGAAATAAGAAACATAATGTGATCCGTTGAAGATCGGAATTTTATATATTACTAT 669
Qy 204 HisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAla 223
Db 670 CACAACCTCATATAAGCAGGCGCCGACTCTGTACTGCAATATATACCTGTGTGCC 729
Qy 224 LysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyrValAsnGly 243
Db 730 AAAACATTTGTTGCCAAGAGGAAAAAGTACAACCTGCCACTGTCTATGTCTCAACGGT 789
Qy 244 GlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGln 263
Db 790 GGTGTGTCCACTGAGCGAGTGTCTGTGTGAACAGCGCTGTGGAGGAGGTATCAG 849
Qy 264 LysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyClyAlaPheCysGluGly 283
Db 850 AAACGTACAGGACTTGTACCAACCCGGCACCACCTCAATGGGGGTGCCCTTCTGTGAAGG 909
Qy 284 GlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrpSer 303
Db 910 CAGAGTGTG---CAGAAATAGCTGTACTACGTATGCTCCAGTGGATGGCAGGTGAGC 966
Qy 304 ProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCys 323
Db 967 CCATGGAGCAAGTGTCTACTTGTGGAAGTGTGACCCACTGGCGCAGGAGGAGTGC 1026
Qy 324 SerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArg 343
Db 1027 ACGGCGCCAGCCCCAGAAATGAGCAAGGACTCGATGGCCTCGTCTTGCAATCCCAAG 1086
Qy 344 AsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363
Db 1087 AACTGCACGTATGGGCTTTGATGATGAGACTGCTCTGATTCAGATGATGTTCTCTAT 1146
Qy 364 ValGlyLeu---IleAlaValAlaValCysLeuValLeuLeuLeuValLeuLeu 382
Db 1147 GTTGGGATGTTGATAGCAGTGTCTTTCCTGGCGATCTCTGTAGTTGGCCCTGTTT 1206
Qy 383 ValTyrCysArgLysLysGluClyLeuAspSerAspValAlaAspSerSerIleLeuThr 402
Db 1207 GTGTATCGGAAGAATCATCGTGACTTTTGAGTCAGATATATTGACTCTTTCGGCACTCAAT 1266

Qy 403 SerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeuThr 422
Db 1267 GGGGGCTTTTCAGCTGTGTGAACATCAAGCAGCAACAAGAT-----CTGTGGCT 1317
Qy 423 IleGlnProAspLeu---SerThrThrThrThrTyrGlnGlySerLeuCysProArgGln 441
Db 1318 GTACCCCGACACCTTCAGTTCAGTGCAGCCATGTACAGAGGACCTGTCTATGCCCTGCAT 1377
Qy 442 AspGlyProSerProLysPheGlnLeuLeuAsnGlyHisLeuLeuSerProLeuGlyGly 461
Db 1378 GAC---GTCTCAGACAAAAATCCCAATGACCAACTCTCCAATTCGGATCCACTGCCCAAC 1434
Qy 462 GlyArgHisThrLeuHisHisSerSer-----ProThrSerGluAlaGluGlu 477
Db 1435 CTGAAAGATCAAGTGTACAACACCTCAGTGTCTGTACCCCCCAAGATGACCTCTCTGAG 1494
Qy 478 PheValSerArgLeuSer-----ThrGlnAsnTyrPhe----- 488
Db 1495 TTTACGTCCAAGCTGTCTCCCTCAGATGCCAGTCACTTGTGGAGAAATGAAGCCCTCAGC 1554
Qy 489 -----ArgSerLeuProArgGlyThr-----SerAsnMetThrTyrGlyThrPhe 503
Db 1555 CTGAAGAACCCAGAGTCTAGCAAGCAGCAGTATCCATCTCTACCGCATTTGGCAGGTTT 1614
Qy 504 AsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIleProPro 523
Db 1615 NNNNGCTGGNNNNNCACTTATTTGTTCCCAATTCAGGAGTCACTGTCTGATTCCCGCT 1674
Qy 524 AspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAsp 543
Db 1675 GGGGCCATTTCCCAAGGAGAGTCTAGAAATGTATGTACTGTACACAGAAAGAAACT 1734
Qy 544 ValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyPro 563
Db 1735 ATGAGGCCACCCATGATGACTCTCAGACACATTTTGCACCCCTGTGTGAGCTGTGGGCC 1794
Qy 564 ProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSer 583
Db 1795 CAGGAGCTCTGTCTACCCGCCCGCTGCTCTTACTATGATCACTGCGCAGACCCCAAT 1854
Qy 584 ProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAsp 603
Db 1855 ACCGAGCATGGAAATATCTGCTCAAGAACCCAGGACGACAGGACAGTGGGAG---GAT 1911
Qy 604 ValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSer 623
Db 1912 GTGGTGTGTGGGAGGAGAAACTTTCACCCACCCCTGTCTACATTCAGCTGGATGCAGAG 1971
Qy 624 AlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSer 643
Db 1972 GCTGCGCACATCTCTCACAGAAACCTCAGCCTACGCCCTGTAGGACATTTCCACCACC 2031
Qy 644 ValAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeu 663
Db 2032 AAAGCGGTGCGAAGCGCTCAAGTGGCCATCTTTGGGCCCTGTGTGCTCTCTCGCTG 2091
Qy 664 GluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValVal 683
Db 2092 GAGTACAGCATCCGAGTCTACTGTCTGGATGACCCAGGATGCCCTGNAGGAAATTTTA 2151
Qy 684 GlnLeuGluLysGlnLeuGlyGlnLeuIleGlnGluProArgValLeuHisPheLys 703
Db 2152 CATCTTGAGACAGATGGAGGACAGCTCTCTAGAAAGACCTAAGGCTCTTCATTTTAA 2211
Qy 704 AspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSer 723
Db 2212 GGCAGCCCAACACCTCGGCTGTCAATTACGATATCGCCCATTCCTCTCTGGAAGAGC 2271
Qy 724 LysLeuLeuValSerTrpGlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArg 743
Db 2272 AAATTGCTGGCTAAATATACGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2331

Direct Submission		Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA	
COMMENT		This sequence was made by sequencing genomic exons and ordering them based on alignment.	
FEATURES		Location/Qualifiers	
source		1..2775	
organism="Homo sapiens"		/mol_type="genomic DNA"	
/db_xref="taxon:9606"		<1..2775	
locus_tag="HCM0901"			
ORIGIN			
Alignment Scores:			
Pred. No.:		Length: 2775	
Score:		Matches: 407	
Percent Similarity:		Conservative: 158	
Best Local Similarity:		Mismatches: 306	
Query Match:		Indels: 52	
DB:		Gaps: 12	
US-09-970-944-2 (1-899) x AY401469 (1-2775)			
Qy	12	LeuGlyIleValLeuAlaThrLeuArgGlySerGlyAlaGlnGlnSerAlaThrVal	31
Db	14	GTAGGCTGTGTTCTGATGCTCTTTTGTGAT-TTCACAGGAAGTGAATGCGAAGCCCTT	72
Qy	32	AlaAsnProValProGlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGlu	51
Db	73	CCCGAATCCATCCCATCAGCT---CCTGGACACTGCTCATTTTCATAGAGGACCAT	129
Qy	52	AspValTyrIleValLysAsnLysProValLeuLeuValCysLysAlaValProAlaThr	71
Db	130	GATGCTTATATATACAGACCAACCTATTTCACCTCAGGTGCAAGCGAGGCCAGCATG	189
Qy	72	GlnIlePhePheLysCysAsnGlyGluTyrValArgGlnValAspHisValIleGluArg	91
Db	190	CAGATATTCTTCAATGCAACGGCGAGTGGGTCCATCAGAACGACGACGCTCTCTGAAGAG	249
Qy	92	SerThrAspGlySerSerGlyGluProThrMetGluValArgIleAsnValSerArgGln	111
Db	250	ACTCTGCAGACAGCTCAGTTTGAAGTCCCGAAGTGTTCATCAATGTTACTAGGCAA	309
Qy	112	GlnValGluLysValPheGlyLeuGluGluTyrTyrCysGlnCysValAlaThrSerSer	131
Db	310	CAGGTGGAGGACTTCCATGGCCCGAGGACTATTGGTGCAGTGTGTGGCTGGAGCCAC	369
Qy	132	SerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPhe	151
Db	370	CTGGGTACTCTCAAGACAGGAGGCGCTCTGTGCGCATAGCCTATTATTACGGAAAACATT	429
Qy	152	GluGlnGluProLeuAlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArg	171
Db	430	GAACAAAGACCCACAGAGGAGGAGTCCCATTTGAAGCATGATTTACTGCACCTCCGC	489
Qy	172	ProProGluGlyIleProProAlaGluValGluTyrLeuArgAsnGluAspLeuValAsp	191
Db	490	CCACCAGAGGAGTCCCTGTGTCGCGAGGTGGAATGGCTGAAAAATGAAGAGCCCATTTGAC	549
Qy	192	ProSerLeuAspProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAla	211
Db	550	TCTGAACAGACGAGAACATTGACACCGAGGCTGACCATTAACCTGATCATCAGGACGCA	609
Qy	212	ArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArg	231
Db	610	CGGCTCTCGGACTCAGAAATTACCTGTCATGCGACCCACATCTGCTGGCTAAGAGGAGA	669
Qy	232	SerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTyrSerThrTyrThrGluTyr	251
Db	670	AGCCTGTGCGCCACTGTTTGGTCTACGTGAATGAGGAGTGGTCTCTCTCGGACAGAGTGG	729
Qy	252	SerValCysSerAlaSerCysGlyArgGlyTyrGlnLysArgSerArgSerCysThrAsn	271

```
Db 1774 CGGACTGTGCAGATGTCTGAGCATTCGAGCATATCCATTTAAAGAGAGACACAG 1833
Qy 597 GluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyr 616
Db 1834 CAGGCAAAATGGAG---GAAGTGTATGTCAGTGGAGATGATCTACATCC-----TGT 1884
Qy 617 TyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAla 636
Db 1885 TACTGCTTTTGGACCCCTTGGCGTGTCTGCTGCTGGACAGCTTTGGGACCTATGCG 1944
Qy 637 LeuValGlyGluAlaLeuSerValAlaAlaAlaLysLeuGlyLeuLeuPheAla 656
Db 1945 CTCACTGAGAGCCAAATCAGACATGTCGCCGTGAAGCAACTGAAGGTGGCGTTTGGC 2004
Qy 657 ProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHis 676
Db 2005 TGCATGCTCTGTAACCTCCCTGGATTACAACTTGAGAGTTTACTGTGGACATACCCCT 2064
Qy 677 AspAlaLeuLysGluValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGlu 696
Db 2065 TGTGCATTTCAGGAAGTGGTTTCAGATCAAGGATCAAGGTGCACAGCTCCTCGAAGAA 2124
Qy 697 ProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspVal 716
Db 2125 CCMAANTGCTGCATTTCAAGGGAATACCTTTAGTCTTCAGATTTCTGCTCTTGATATT 2184
Qy 717 ProSerSerLeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHis 736
Db 2185 CCCCATTCTCTGGAGATTAAACCATTCAGTCCCTGCAGGAAGTCCCGTCTCCCGC 2244
Qy 737 IleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerPro 756
Db 2245 GTGTGTGCAGTAAACCGGAGCCCTGCATGTGCTTCTCCCTGGAGCGTTATACGCC 2304
Qy 757 SerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSer 776
Db 2305 ACTACCAACCCAGCTGCTGCAAAATCTGCAATTCGGCAGCTCAAGGCCATGAACATC 2364
Qy 777 PheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuAlaLeuGlu 796
Db 2365 CTCCAAGTGCAGATCAATCTTAGAGAGTGAACGAAACCATCACTTTCTTCGCACAA 2424
Qy 797 SerGluAlaGlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIle 816
Db 2425 GAGGACAGCATTTTCCTGTCACAGACTGGCCCAAGCCCTTCAAAATTCCTACTCCATC 2484
Qy 817 ArgGlnLysIleIleSerSerLeuAspProCysArgArgGlyAlaAspTrpArgThr 836
Db 2485 AGACAGCGATTGTGCTACATTTGATACCCCAATGCCAAGGCCAAGGACTGGCAGATG 2544
Qy 837 LeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerPro 856
Db 2545 TTAGCACAGAAACACATCAACAGGAATTTATCTTATTTGCTACAAAGTAGCCCA 2604
Qy 857 ThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGln 876
Db 2605 TCTGCTGTCAATTTGAACCTGTGGAGCTCGTCATCAGCATGATGTGTGATCTTGACTCC 2664
Qy 877 LeuAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGlu 896
Db 2665 CTGGCTGTGCCCTTGAAGAGATTGGGAGGACACACACGAAAACCTCTCAAAACATTTAGAA 2724
Qy 897 AlaGlu 898
Db 2725 TCCCAG 2730

RESULT 10
AY411748
LOCUS
DEFINITION
Pan troglodytes HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
AY411748
ACCESSION
AY411748.1 GI:39767716
```

```
KEYWORDS
SOURCE
ORGANISM
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 (bases 1 to 2532)
AUTHORS
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
REFERENCE
2 (bases 1 to 2532)
AUTHORS
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE
Direct Submission
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source
1..2532
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>2532
/locus_tag="HCM4327"
gene
ORIGIN
Alignment Scores:
Pred. No.: 3,76e-157 Length: 2532
Score: 1890.00 Matches: 390
Percent Similarity: 60.87% Conservative: 125
Best Local Similarity: 46.10% Mismatches: 275
Query Match: 39.48% Indels: 56
DB: 29 Gaps: 16
US-09-970-944-2 (1-899) x AY411748 (1-2532)
Qy 103 GluValArgIleAsnValSerArgGlnValGluLysValPheGlyLeuGluGluTyr 122
Db 13 GAGGTGCAGATCGAGGTGTGGCGCAGCAGGTGGAGAGCTCTTTGGGTGGAGGATTAC 72
Qy 123 TrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAlaTyrIle 142
Db 73 TGGTGCAGTGTGGTGGCTGGAGCTCTGGGGCACCACCAAGAGTCGCGAGCCTACGTC 132
Qy 143 ArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeu 162
Db 133 CGCATCGCTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 192
Qy 163 GluGlnGlyIleValLeuProCysArgProGluGlyIleProAlaGluValGlu 182
Db 193 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 252
Qy 183 TrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArg 202
Db 253 TGGCTCAAGATGAGGATGTCTATCGACCCCAAGGAGGAGGAGGAGGAGGAGGAGGAG 312
Qy 203 GluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysVal 222
Db 313 GACCAACACCTCATCTCGCGAGCGCCCTGTCTGGACACTGCCAATCTATACCTGCGTG 372
Qy 223 AlaLysAsnIleValAlaArgArgSerAlaSerAlaValIleValTyrValAsn 242
Db 373 GCCAAGAACATCGTGGCCAAACGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 432
Qy 243 GlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrp 262
```

Db	433	GGCGGCTGCTCCAGCTGGCGCAGAGTGGTGGCCCTGCTCCAAACCGCTGTGNNNNCTGG	492	576	MetAspHisCysGlyGluProSerProAspSerTyrSerLeuArgLeuLysLysGlnSer	595
Qy	263	GlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGlu	282	1558	ATGCCCATCTGTGCCCAAGTCACTGCTCCCTGCTGAGTCTTTCAGTCAAGACCCAGGCC	1617
Db	493	CAGACGGCAGCCGACCTGACCAACACCCGCCCTCAACGAGGNNNNCTCGGAG	552	596	CysGluGlySerTyrGlnGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeu	615
Qy	283	GlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTyr	302	1618	CACCAGGCCACCTGGGAG--GAGGTGGTGACCTCGATGAGGAGACCTGAAACACACC	1674
Db	553	GGC---CANNN	609	616	TyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPhe	635
Qy	303	SerProTyrSerLysTyrSerAlaCysGlyLeuAspCysThrHisTyrArgSerArgGlu	322	1675	TGCTACTGCCAGCTGAGGCCCGCTGTCACTCTGCTGACACCTGCGGACCTGACCTAC	1734
Db	610	ACCGAGTGGAGCAAGTGGTCACTGAGTGGTCCCGAGTGGTCCCGAGTGGTCCCGAG	669	636	AlaLeuValGlyGluAlaLeuSerValAlaAlaAlaAlaArgLeuLysLeuLeuPhe	655
Qy	323	CysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThr	342	1735	GTGTTCCAGGGCGAGTCTTATTCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1794
Db	670	TGATGCGGCCCTCAGG---GATGCGCGCTGTATGCGGGCTCGTGGTGCC	729	656	AlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThr	675
Qy	343	ArgAsnCysThrSerAspLeuCysValHisSer-----	353	1795	GCCCTGCTGCTGACCTCCCTGGANNNNNGCTCCGGTCTACTGCTGNNNNNNNN	1854
Db	730	AAGAACTGCACAGATGGCTGTGATGCAAAATAAGAAAACTCTAAGCGACCCCAACAGC	789	676	HisAspAlaLeuLysGluValValGlnLeuGlyGlnLeuGlyGlnLeuGln	695
Qy	354	-----AlaSerGlyProGluAspValAlaLeuTyrValGlyLeu---IleAla	368	1855	NN	1914
Db	790	CACCTGCTGAGGCTCAGG---GATGCGCGCTGTATGCGGGCTCGTGGTGCC	843	696	GluProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAsp	715
Qy	369	ValAlaValCysLeuValLeuLeuValLeuValLeuValTyrCysArgLysLys	388	1915	GAGCGGAAACCGTATGTTCAAGGACAGTTCACCAACCTGCGCTCTCCCTCCATGAC	1974
Db	844	GTCTTCNTGTCNTGGCAATCTCATGGCGTGGGGTGGTGGTACCGCGCACTGC	903	716	ValProSerSerLeuTyrLysSerLysLeuValSerTyrGlnGluLeuProPheTyr	735
Qy	389	GluGlyLeuAspSerAspValAlaAspSerSer---IleLeuThrSerGlyPheGlnPro	407	1975	CTCCCCATGCCATTGGAGGAGCAAGCTGCTGGCCAAATACCAAGGAGATCCCTCTGT	2034
Db	904	CGTGACTGCACACAGACATCACTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	963	736	HisIleTyrAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSer	755
Qy	408	ValSerIleLysProSerLysAlaAspAsnProHisLeuLeu-----ThrIleGlnPro	425	2035	CACATTGGAGTGGCAGCAGAGCCCTCCACTGCACCTTTCACCTCGAGAGGACACGC	2094
Db	964	GTCAACTTTAAGCGGCAAGCCCGACCAACCGGAGCTCTTACACCCCTCTGCTGCTCT	1023	756	ProSerThrSerAspLeuAlaCysLysLeuTyrValTyrGlnValGluGlyAspGlyGln	775
Qy	426	AspLeuSerThrThrThr---ThrTyrGlnGlySerLysCysProArgGlnAspGlyPro	444	2095	TGGCTCTCCACAGCTACCTGCAAGATCTGCGTGGCGCAAGTGGAGGAGGAGGCCAG	2154
Db	1024	GACTGACAGCAGCGCGCGGCTATACCGGAGCCGCTGATGCCCTGCTGCTGCTGCTGCT	1080	776	SerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeu	795
Qy	445	SerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyArgHis	464	2155	ATATTCCAGCTGCATACCACTCTGGCA---GAGACACCTGCTGCTGCTGCTGCTGCT	2211
Db	1081	ACCAGCAAAATCCCATGACCACTCTCTCTGCTGGACCCCTTACCCAGCCTTAAAGTC	1140	796	GluSerGluAlaGly-----ValProAlaLeuValGlyProSerAlaPheLysIlePro	813
Qy	465	ThrLeuHisHisSerSerProThr-----SerGluAlaGluGluPhe	478	2212	TGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2271
Db	1141	AAGTCTTACAGCTCCAGCACTACCGGCTCTGGGCGAGCTGGGAGTGGGCTGACCTG	1200	814	PheLeuIleArgGlnLysIleSerSerLeuAspProProCysArgArgGlyAlaAsp	833
Qy	479	ValSerArgLeuSerThrGlnAsnTyr-----	487	2272	CTGTCCATCCCGCAGAGTATGCAACAGCTAGATGCCCACTACCGGGGCAATGAC	2331
Db	1201	CTGGAGCTTGTCCCGCTGGCAGCATACCTAGCGATTTTCGCCGGGACACCCACTTCCT	1260	834	TyrArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLys	853
Qy	488	---PheArgSer-----	496	2332	TGGCGGATGTAGCACAGAGCTCTCTATGACCGGTACCTGAAATTTCTTGTCCACCAA	2391
Db	1261	CACCTGCGCAGCGCGCAGCTCGGTTCCCGAGCAGCTCTTGGGCTCGCCCGAGACCCAGN	1320	854	ProSerProThrAlaMetIleLeuAsnLeuTyrGluAlaArgHisPheProAsnGlyAsn	873
Qy	497	AsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGly	516	2392	GCGAGCCCCACGGGTGTCTCTGACCTCTGGGAGCTCTGACAGCAGAGATGGGAC	2451
Db	1321	AGCAGCGTCAGCGGCACCTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1380	874	LeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThr	893
Qy	517	IleSerLeuLeuIleProProAspAlaIleProArgGlyIleIleTyrGluIleTyrLeu	536	2452	CTCAACAGCTTGGCGAGTGGCTTGGAGGAGATGGCGAGAGTGGATGCTGTGGTGTGTG	2511
Db	1381	GTGAGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1440	894	ValSerGluAlaGluCys	899
Qy	537	ThrLeuHisLysProGluAspValArgLeuProLeuAla---GlyCysGlnThrLeuLeu	555	2512	GCCACCGGCGGACTGC	2529
Db	1441	CTCATCAACAGCGAGAAAGTAGTACC---CTCCGCTTTTCAAGAGGGACCCAGACATG	1497	RESULT 11		
Qy	556	SerProIleValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAla	575	BC033727		
Db	1498	AGCCCTCTGAGTACCTGTGAGCCACAGGCTCTGCTGTGCAACCCCTGCTCCTCACC	1557	LOCUS		
				DEFINITION	Homo sapiens netrin receptor Unc5h1, mRNA (cDNA clone	
					BC033727	
					1532 bp	
					mRNA	
					linear	
					HTC	
					19-NOV-2003	

IMAGE:5166762), containing frame-shift errors.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BC033727
BC033727.1 GI:21707230
HTC.
Homo sapiens (human)

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
 1 (bases 1 to 1532)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Moore, T.K., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T.K., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Spatacchio, M., Soares, M.B., Bonaldo, M.F., Cabavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Rulyk, S.W., Vallalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Rodrigues, S., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schermer, A., Schein, J.E., Jones, S.J. and Marra, M.A.,
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 223388257
 2 (bases 1 to 1532)
 Strausberg, R.
 Direct Submission
 Submitted (02-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgaps@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Gaithersburg, Maryland,
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nih.gov
 Khokhr, N., Aylee, K., Beckstrom-Sternberg, S.M., Benjamin, B., Kistler, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S., Kistler, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, F., Legasi, R., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stanterop, S., Thomas, P.J., Touchman, J.W., Voigt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK plate: 68 Row: i Column: 2
This clone was selected for full length sequencing because it passed the following selection criteria: GenomScan gene prediction
This clone has the following problem: frame shifted.

FEATURES	SOURCE
----------	--------

```
Location
1. .1532
/organism
/mol_type
/db_xref:
/clone="1"
/tissue
/clone 1:
```

```
/note="vector: pCMV-SPORT6"
```

/note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores:

Pred. No.:	2,91e-155	Length:	1532
Score:	1864.50	Matches:	377
Percent Similarity:	67.08%	Conservative:	2
Best Local Similarity:	66.73%	Mismatches:	184
Query Match:	38.95%	Indels:	3
DB:	11	Gaps:	

US-09-970-944-2 (1-899) x BC033727 (1-1532)

23	SerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnProAspLeu	42
Db	GCAGGTGCCACGACAGAGTGCACCGTGGCCAAACCCAGTGCCTGGTGCCAAACCCGGACCTG	70
43	LeuProHisPheLeuValGluProGluAspValTyrIleValIysAsnLysProValLeu	62
Db	CTTCCCCACTTCCCTGGTGGAGCCGAGGATGTGTACATCTCAAGAACCAAGCAGCAGTGCTG	130
63	LeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpVal	82
Db	CTTGTGTGCAAGGCCGTGCCGCCCAACCCAGATCTTCTCAAGTGCACCGGGAGTGGGTG	190
83	ArgGlnValAspHisValIleGluArgSerThrAspGlySerSer	97
Db	CGCCAGGTGGACCAACGTCATCGAGCGCAGCACAGACGGGAGCAGTGAGGGAGATAGAGAG	250
97	-----	97
251	GGTTGGCGGCAATGTAGCCAAACAGAACCCCTGGGCACCAACACCTCCCGAGGACCAGAGC	310
97	-----	97
311	CTCTCCCGCCCCAGCGGAGGTCCAGAGGCCCTCCTGCCCGACGACGAGGGGCCACCTTG	370
97	-----	97
371	TC TTGCTGGAGGGTCTGCAGGAGAACCGCAGGGTCCTCTTCTGCCCGCAGGACATCGACT	430
97	-----	97
431	CCAGCCTCCCATGGCCCGGAACATCTGAACGTTCAITGATCTCATCGATTTCTCAACCCAA	490
97	-----	97
491	AGCCATCGAGTGCTTTGAGGTGAAGAAAGGCTTTCCTTACCCAGCGAGATACCCAG	550
97	-----	97
551	CAGTGGCGCAACGCGCCCAAGACCAAGACCCCAACCAAGAGCTTCTGTGTCAGAC	610
98	-GlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValph	117
611	AGGGCTGCCCCACATGGAGGTCCGATTAATGTCTCAAGGCAGCAGGTCCGAGAAGTGTT	670
117	eGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSe	137
671	CGGGCTGGAGGAATACTGTGTCCAGTGGTGATGGAGTCTCTCGGGCACCAACAGAG	730
137	rGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAl	157
731	TCAGAGGCTACATCCGCATAGCCTATTTGGGCAGAACTTTCAGACGAGGAGCCGCTGGC	790
157	alysGluValSerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIlePr	177
791	CAAGGAGGTGTCCCTGGAGCAGGGCATCGTCTGCCTCGCTCCACCGGAGGSCATCCC	850
177	oProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAs	197
851	TCCAGCCGAGGTGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCTTGGACCCCAA	910

QY 197 nValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAl 217
Db 911 TGTATACATCAGCGGAGCACAGCCTGTGTGTGCGACAGCGCCGCTTGTGACACGGC 970
QY 217 aAntyThrCysValAlaLysAsnIleValAlaArgArgAcqSerAlaSerAlaVala 237
Db 971 CAACACTACCTGGTGGCCAAAGAACATCGTGGCACGTGGCGGACGCGCTTCCCTGCTGT 1030
QY 237 lIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSe 257
Db 1031 CATCGTCTAC----- 1040
QY 257 rCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGl 277
Db 1040 ----- 1040
QY 277 yGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSe 297
Db 1040 ----- 1040
QY 297 rValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHi 317
Db 1041 -GTGGACCGCAGTGTGAGCCCGTGGAGAAAGTGTGGCGCTGTGGGTGACCTGCACCCA 1099
QY 317 sTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGl 337
Db 1100 CTGGCGGAGCGGTGAGTGTCTGTGACCCAGCACCCCGCAAGAGGAGGAGGTGCCAGGG 1159
QY 337 yThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyPr 357
Db 1160 CACTGACCTGGACACCCGCACTGTACAGTGTGCTGTGTGTACACTGTCTTCTGGCC 1219
QY 357 oGluAspValAlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLe 377
Db 1220 TGAGGAGTGGCCCTCTATGTGGCCCTCATCGCGCTGCTGCTGCTGCTGCTGCTGCT 1279
QY 377 uLeuValLeuLeuValTyrCysArgLysGluGlyLeuAspSerAspValAlaAs 397
Db 1280 GCTTGTCTCATCTCTGTTATTGCCGAAGAAGAGGAGGCTGGACTCAGATGTGGCTGA 1339
QY 397 pSerSerIleLeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAs 417
Db 1340 CTCGTCAATTCTACCTCAGGCTTCCAGCCCGTTCAGCATCAAGCCAGCAAGACAGCAA 1399
QY 417 nProHisLeuLeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySe 436
Db 1400 CCCCCATCTGCTCACCATCCAGCGGACCTCAGCACACCACCCACCTACCGGGCAG 1459
QY 436 rLeuCysProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLe 456
Db 1460 TCTCTGTCGGGAGGATGGGCCAGGCCCAAGTTCAGGCTCACCATGGGCACCTGCT 1519
QY 456 uSerProLeuGly 460
Db 1520 CAGCCCCCTGGGT 1532
RESULT 12
AY401470 2507 bp DNA linear GSS 12-DEC-2003
LOCUS AY401470
DEFINITION Pan troglodytes HCM0901 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY401470
VERSION AY401470.1 GI:39757459
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1 (bases 1 to 2507)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D., and Cargill, M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2507)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D., and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..2507
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
gene <1..>2507
/locus_tag="HCM0901"
ORIGIN
Alignment Scores: 9.18e-127 Length: 2507
Pred. No.: 1550.00 Matches: 332
Score: 52.59% Conservative: 135
Percent Similarity: 37.39% Mismatches: 257
Best Local Similarity: 32.38% Indels: 164
Query Match: 29 Gaps: 14
DB: 14
US-09-970-944-2 (1-899) x AY401470 (1-2507)
QY 103 GluValArgIleAsnValSerArgGlnGlnValGlnLysValPheGlyLeuGluGluTyr 122
Db 15 GAAGTGTTCATCATGTATTACAGGCAACAGTGTGAGGACTTCCATGGGCCCGGAGACTAT 74
QY 123 TrpCysGlnCysValAlaTtpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIle 142
Db 75 TGGTGGCAGTGTGTGGGTGGAGCCACTGGTACCTCCAGAGCAGGAGGAGGCTCTGTG 134
QY 143 ArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeu 162
Db 135 CGCATAGCTATTATTACGGAAAACTTTGAAACAGACCCACCAAGGAAGGAGGTTCCCAT 194
QY 163 GluGlnGlyIleValLeuProCysArgProGluGlyIleProProAlaGluValGlu 182
Db 195 GAAGGCATGATTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 254
QY 183 TrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArg 202
Db 255 TGGCTGAAAAATGAAGACCCCATTTGACTCTGAACAGACGAGAACATTGACACAGGGCT 314
QY 203 GluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysVal 222
Db 315 GACCAATACTGATCATCAGGCAGGACGCTCTCGGACTCAGGAAATATACACCTGCATG 374
QY 223 AlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyrValAsn 242
Db 375 GCAGCAACATCTGTGGCTTAAGAGGAGAGCCCTGTGGCCACTGTTGTGTGTACGTGAAT 434
QY 243 GlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrp 262
Db 435 GGAGGCTGCTTCTCTCGACAGAGTGGTTCAGCTGCAATGTTGCTGTGTGTAGAGGATGG 494
QY 263 GlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGlu 282
Db 495 CAGAAACGTTCCGGACCTGCACCAACCCAGCTCTCTCAATGGTGGGCGCTTTGTGTAG 554
QY 283 GlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrp 302
Db 555 GGAATGTCTAGTCAGAAATAATAC----- 578


```
QY 303 SerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGlu 322
Db 578 -----
QY 323 CysSerAspProAlaProArgAsnGlyGlyGluCysGlnGlyThrAspLeuAspThr 342
Db 578 -----
QY 343 ArgAsnCysThrSerAspLeuCys----- 350
Db 579 -----TGCACCTCT-----CITTGCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 629
QY 350 ----- 350
Db 630 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 689
QY 350 ----- 350
Db 690 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 749
QY 351 -----ValHisSerAlaSerGlyProGluAspValAlaLeuTyrValGly 365
Db 750 NNNNNNNNNNNNGCATTTGAGATGCCAGC-----GACATTGCTTTGTACTCGGC 800
QY 366 LeuLeuAlaValAlaValCysLeuValLeuLeuLeuValLeuLeuValLeuValTyrCys 385
Db 801 TTG---GGCGCTCGCGTGGCGGTGCAGTCCTGTCATTGTCATTGTCACCCCTTACAGA 857
QY 386 ArgLysGlyGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSerGlyPhe 405
Db 858 CGGAGCCAGAGTAGTACTATGCGTGGAGCTATTGACTCTTCTGCAATTGACAGGTGGCTTC 917
QY 406 GlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu-----ThrIle 423
Db 918 CAGACCTTCAACTTCAAAACAGTCCGCTCAAGNNNACTCCCTGCTCTGTAATTCGGCCATG 977
QY 424 GlnProAspLeuSerThrThrThrThrTyrGlnGlnSerLysCysProArgGlnAspGly 443
Db 978 CAGCCAGATCTGACAGTGGCCGGACATACAGCGGACCCATCTGT---CTGCAAGGAC--- 1031
QY 444 ProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGly----- 460
Db 1032 CCTCTGGACAAGGAGCTCATGACAGAGTCTCTACTGTTTAAACCTTTGTGCGACATCAMA 1091
QY 460 ----- 460
Db 1092 GTCAAAAGTCCAGAGCTGTTTCATGGTTTCCCTGGGAGTGTCTGAGAGAGCTGAGTACCAC 1151
QY 461 -----GlyGlyArgHisThrLeuHisHisSerSer 470
Db 1152 GGCAAGAATCATTCAGGACTTTTCCCATGGAAACAACACAGCTTTAGTACAATGCAT 1211
QY 471 ProThrSerGluAlaGluGluPheValSerArgLeuSerThrGlnAsnTyrPheArgSer 490
Db 1212 CCCAGAAATAAAATG---CCCTACATCCAAAATCTGTCA-----TCA 1250
QY 491 LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeu 510
Db 1251 CTCCCCACAGGACAGAACTGAGGACAACTGTGTGTCTNNNGCCATTTAGGGGGCGGNNNN 1310
QY 511 MetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLys 530
Db 1311 NNNNNN---NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1367
QY 531 IleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAlaGly 550
Db 1368 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1427
QY 551 CysGlnThrLeuLeuSerProIleValSerCysGlyProGlyValLeuThrArg 570
Db 1428 TCTGAGGTCTCTCTGAGTCTCTGAAGTCACTGTGTGTCTCTCCAGACATGATCTCACCATT 1487
```

```
QY 571 ProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeuArg 590
Db 1488 CCCTTTGCATTGACCATCCACACACTGTGCAGATGTCTAGTNNNGAGCACTGGAAATATCCAT 1547
QY 591 LeuLysGlnSerCysGlyGlySerTrpGluGlnAspValLeuHisLeuGlyGluGlu 610
Db 1548 TTAAGAAGAGGACACAGCAGCGCAATGGAG---GAAGTGAATGTGAGTGAAGATGAA 1604
QY 611 AlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGlu 630
Db 1605 TCTACATCC-----TGTTACTGCTTTTGGACCCCTTTGTCATGTCTGCTCTGGAC 1658
QY 631 GlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaLysArgLeu 650
Db 1659 AGCTTTGGGACCATATCGCTCACTGGAGAGCCAAATCACAGACTGTGCGGTGAAGCAACTG 1718
QY 651 LysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyr 670
Db 1719 AAGTGGCGGTTTTTGGCTGCATGTCTGTAACTCCCTGGATTACAACTTGAGAGTTTAC 1778
QY 671 CysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGlyLysGlnLeuGly 690
Db 1779 TGTGTGGCAATACCCCTTGTGCATTTCCANGAAGTGGTTTCAGATGAAAGGCATCAAGGT 1838
QY 691 GlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeuArg 710
Db 1839 GGACAGCTCTTGAAAGAACCAAAATTTGCTGCATTTCAAGGAATACCTTTAGTCTTCAG 1898
QY 711 LeuSerIleHisAspValProSerSerLeuTyrLysSerLysLeuLeuValSerTyrGln 730
Db 1899 ATTTCCTCTCTGATATTTCCCATTCCTCTGGAGAAATTAACCATTCCTACTGCTGCCAG 1958
QY 731 GluLeuProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThr 750
Db 1959 GAAGTCCCTTCTCCCGCTGTGGTGCAGTAACCCGACGCCCTGCTGCTGCTGCTCTCC 2018
QY 751 LeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnVal 770
Db 2019 CTGGAGCGTATACGCCCACTACCCAGCTGTCTCTGCAAAATCTGCATTCGGCAGCTC 2078
QY 771 GlyGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAla 790
Db 2079 AAAGGCCATGAACAGATCTCTCAAGTGCAGACATCAATCTAGAGAGTGAACAGAAACC 2138
QY 791 GluLeuLeuAlaLeuGluSerGluAlaGlyValProAlaLeuValGlyProSerAlaPhe 810
Db 2139 ATCACTTTCTTTGCAACAGAGGACAGCACTTTCCCTGTCACAGACTGGCCCCAAGCCTTC 2198
QY 811 LysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProProCysArgArg 830
Db 2199 AAAATTCCCTACTCCATCAGACAGCGGATTTGTGTACATTTGATACCCCAATGCCAAA 2258
QY 831 GlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePhe 850
Db 2259 GCAAGAGCTGCACATGTTAGCACAGAAAACAGCATCAACAGGAATTTATCTATTTC 2318
QY 851 AlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPhePro 870
Db 2319 GCTACACAAAGTAGCCCATCTGCTCTATTTTGAACCTGTGGGAAGCTGTCATCAGCAT 2378
QY 871 AsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnProAspAlaGly 890
Db 2379 GATGTGATCTTGAATCTCCCTGCTGCTGCTGCTGCTTGAAGAGATTGGGAGGACACACGAAA 2438
QY 891 LeuPheThrValSerGluAlaGlu 898
Db 2439 CTCTCAACATTTCAATATCCAG 2462
```

RESULT 13
BI758231
LOCUS
DEFINITION
603029876F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200171 5',
mRNA sequence.

```

ACCESSION BI758231
VERSION BI758231.1 GI:15749809
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1034)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11501 row: 9 column: 20
High quality sequence stop: 793.
Location/Qualifiers
1..1034
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5200171"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 1,13e-111 Length: 1034
Score: 1373.50 Matches: 291
Percent Similarity: 88.82% Conservative: 11
Best Local Similarity: 85.59% Mismatches: 28
Query Match: 28.69% Indels: 14
DB: 12 Gaps: 2

US-09-970-944-2 (1-899) x BI758231 (1-1034)

QY 36 ProGlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrIle 55
Db 2 CCTGGTCCACCCGGACCTCTTCCCACTTCTCTGGTGGAGCCGAGGATGTGTACATC 61
QY 56 ValLysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePhe 75
Db 62 GTCAAGAACCAAGCCAGTGTCTTGTGTGCAAGCCGTCGCCGCCAGCAGATCTTCTTC 121
QY 76 LysCysAsnGlyGluTTPValArgGlnValAspHisValIleGluArgSerThrAspGly 95
Db 122 AAGTGCACCGGGAGTGGTGGCCAGGTGGACCAACACATGATCGAGCGCAGCAGACGGG 181
QY 96 SerSerGlyGluProThrMetGluValArgIleAsnValSerArgGlnValGluLys 115
Db 182 AGCAGTGGCTGCCACCATGGAGTCCGCATTAATGTCTCAAGGCAGCAGGTGCAGAG 241
QY 116 ValPheGlyLeuGluGluTyrTTPCysGlnCysValAlaTTPSerSerGlyThrThr 135
Db 242 GTGTTCGGGCTGGAGGATATCTGGTGCAGTGGGATGAGTCTCTCGGGACCAACC 301
QY 136 LysSerGlnLysAlaTyrIleArgIleAlaArgLeuLeuArgLysAsnPheGluGlnPro 155

```

```

Db 302 AAGAGTCAGAGGCGCTACATCCGCATAGCTATTTTCGCCAGAACTTCGAGCAGGAGCG 361
QY 156 LeuAlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProGluGly 175
Db 362 CTGGCCAGGAGGTGTCCCTCGAGCAGGCACTCGTGTCCCTCGCTCCACCGAGGCG 421
QY 176 IleProProAlaGluValGluTTPLeuArgAsnGluAspLeuValAspProSerLeuAsp 195
Db 422 ATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCTGGAC 481
QY 196 ProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAsp 215
Db 482 CCCAATGTATACATCACGCGGAGCACAGCTGTGTGTCGACAGGCGCCCTTGTGAC 541
QY 216 ThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgSerAlaSerAla 235
Db 542 ACGGACAACATACACCTGCGTGGCCAGAACATCTGTGGCAGCTGCCGCGAGCGCTCGCT 601
QY 236 AlaValIleValTyrValAsnGlyGlyTTPSerThr-TTPThrGluTTPSerValCysSe 255
Db 602 GCTGTCAATGCTTACGTGAACGGTGGTGGTGGACCGTGGACCGAGTGGTCCGCTGCGC 661
QY 255 rAlaSerCysGlyArgGlyTTPGlnLys-ArgSerArgSerCysThrAsnProAlaProL 275
Db 662 -GCCAGCTGTGGCGCGGCTGGCAGAAACCGGAGCGGAGCTGCAC-AAACCCGCTGCTC 719
QY 275 euAsnGlyGlyAlaPheCysGlu-GlyGlnAsnValHisAspArgThrValSerSerLeu 294
Db 720 TCAACGGGGCGCTTCTGTGAGGGGCGAATGTC-----CAGAAAGCAGCTGCCCA 773
QY 295 LeuValSerValAspGlySerTTPSerProTTPSerLysTTPSerAlaCys--GlyLeuA 314
Db 774 CCTGTGCCAGTGAGCGGAGCTGTAGCACGTGGAGCCAGTGGTGGCTGTGGGCTGG 833
QY 314 spCysThrHis-TTPArgSerArgGluCysSerAsp-ProAlaProArgAsnGlyGlyG 333
Db 834 CTTTGACCCCACTTGGCGGAGCGGAGTGTCTGAAACCCAGACCCCGGACAGGGG 893
QY 333 uGluCys--GlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHi 352
Db 894 GAGGTGTGCCCGCTACCTGGACCTGGACCCCGGGATGG-TCCAGTGGAGTCTGTGTCCC 952
QY 352 sSerAlaSerGlyProGluAspValAlaLeuTyrValGlyLeuIle 367
Db 953 -ACGGGTCTTGGCCCTGAGGACTTGGCTCTCTAATGTGGGCTCATC 997
RESULT 14
BI818609 818 bp mRNA linear EST 04-OCT-2001
LOCUS 603033362F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174559 5',
DEFINITION mRNA sequence.
ACCESSION BI818609
VERSION BI818609.1 GI:15929902
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 818)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11434 row: 1 column: 16
High quality sequence stop: 744.

```


Score: 1274.00 Matches: 255
Percent Similarity: 63.59% Conservative: 0
Best Local Similarity: 63.59% Mismatches: 0
Query Match: 26.61% Indels: 146
DB: 13 Gaps: 1

US-09-970-944-2 (1-899) x BQ689148 (1-874)
QY 499 ThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSer 518
Db 1 ACCTATGGGACCTTCACTTCCTCGGGGCGGCTGATGATCCCTAATACAGGAATCAGC 60
QY 519 LeuLeuIleProProAspAlaIleProArgGlyIleTyrGluIleTyrLeuThrLeu 538
Db 61 CTCCTCATCCCCCAGATGCATACCCCGAGGGAAGATCTATGAGATCTACCTACCGCTG 120
QY 539 HisIleProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIle 558
Db 121 CACAAGCCGAAGACGTGAGGTTGCCCTAGCTGGCTGTCTCAGACC----- 165
QY 559 ValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHis 578
Db 165 ----- 165
QY 579 CysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGly 598
Db 165 ----- 165
QY 599 SerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCys 618
Db 165 ----- 165
QY 619 GlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuVal 638
Db 165 ----- 165
QY 639 GlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProVal 658
Db 165 ----- 165
QY 659 AlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAla 678
Db 165 ----- 165
QY 679 LeuLysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnProArg 698
Db 165 ----- 165
QY 699 ValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSer 718
Db 166 ---CTGCACCTTCAAGGACAGTTACCACCACTGGCGCTTATCCATCCAGATGGCCAGC 222
QY 719 SerLeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrp 738
Db 223 TCCCTGTGGAAGAGTAAGTCTCTGTGCTACCTACAGGAGATCCCTTTTATCAGATCTGG 282
QY 739 AsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThr 758
Db 283 AATGGCAGCGCGGTACTTGCACCTGCACCTTACCCCTGGAGCGTGTACGCCCCAGCACT 342
QY 759 SerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSer 778
Db 343 AGTGACCTGGGCTGCAAGCTGTGGGTGGGAGGTGGAGGGGCGAGCGGAGCTTCAGC 402
QY 779 IleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuAlaLeuGluSerGlu 798
Db 403 ATCACTTCAACATCACCAAGACACAGGTTTGTGCTAGCTGCTGCTGGAGAGTAA 462
QY 799 AlaGlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGln 818
Db 463 GCGGGGGTCCAGCCCTGGTGGGCCCCAGTGTCTTCAAGATCCCTTCTCATTCGGGAG 522
QY 819 LysIleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAla 838

Search completed: October 6, 2004, 22:37:02
Job time : 5582 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2004, 09:39:46 ; Search time 11136 seconds
(without alignments)
11213.297 Million cell updates/sec

Title: US-09-970-944-1

Perfect score: 2881

Sequence: 1 agctggggtccgggtgag.....ccttcccccacacggggaga 2881

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb.ba:*

2: gb.htg:*

3: gb.in:*

4: gb.om:*

5: gb.ov:*

6: gb.pat:*

7: gb.ph:*

8: gb.pl:*

9: gb.pr:*

10: gb.ro:*

11: gb.sts:*

12: gb.sy:*

13: gb.un:*

14: gb.vi:*

15: em.ba:*

16: em.fun:*

17: em.hum:*

18: em.in:*

19: em.mu:*

20: em.om:*

21: em.or:*

22: em.ov:*

23: em.pat:*

24: em.ph:*

25: em.pl:*

26: em.ro:*

27: em.sts:*

28: em.un:*

29: em.vi:*

30: em.htg.hum:*

31: em.htg.inv:*

32: em.htg.other:*

33: em.htg.mus:*

34: em.htg.pln:*

35: em.htg.rod:*

36: em.htg.mam:*

37: em.htg.vrt:*

38: em.sy:*

39: em.htgo.hum:*

40: em.htgo.mus:*

41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2881	100.0	2881	6	AX527916	AX527916 Sequence
2	2676.4	92.9	2752	6	AX449572	AX449572 Sequence
3	2615	90.8	2697	6	AX451652	AX451652 Sequence
4	2402.6	83.4	3580	6	AX367094	AX367094 Sequence
5	2327.8	80.8	3992	10	MMU487852	MMU487852 Sequence
6	2235.8	77.6	3014	6	BD057524	BD057524 Netrin re
7	2189.4	76.0	2697	6	AX268596	AX268596 Sequence
8	2189.4	76.0	2697	10	RNU87305	RNU87305 Sequence
9	2020.6	70.1	3844	10	BC058084	BC058084 Mus muscu
10	1691.4	58.7	2688	9	BC009333	BC009333 Homo sapi
11	1570.4	54.5	1787	6	BD057525	BD057525 Netrin re
12	1291.2	44.8	9700	6	AX054976	AX054976 Sequence
13	940.2	32.6	9299	10	MMU72634	MMU72634 Sequence
14	938.4	32.6	2962	5	AY187310	AY187310 Gallus ga
15	936.8	32.5	9328	10	AB118026	AB118026 Rattus no
16	916	31.8	3646	9	AF055634	AF055634 Homo sapi
17	895	31.1	3770	9	AY126437	AY126437 Homo sapi
18	889.4	30.9	4294	10	AK122575	AK122575 Mus muscu
19	889	30.9	2860	6	AX686445	AX686445 Sequence
20	887.4	30.8	2860	6	AX686447	AX686447 Sequence
21	875.8	30.4	2995	6	AX497288	AX497288 Sequence
22	866.4	30.1	2895	6	AX512281	AX512281 Sequence
23	859.8	29.8	3672	10	BC057560	BC057560 Mus muscu
24	859.8	29.8	3672	10	BC048162	BC048162 Mus muscu
25	856.8	29.7	3933	9	AK128132	AK128132 Homo sapi
26	856.8	29.7	3935	9	AB096256	AB096256 Homo sapi
27	853.6	29.6	3884	6	AX464012	AX464012 Sequence
28	853.6	29.6	3884	9	AY358351	AY358351 Homo sapi
29	821.6	28.5	3788	10	MMU487853	MMU487853 Mus muscu
30	794.4	27.6	2838	10	RNU87306	RNU87306 Rattus norv
31	783.4	27.2	2831	6	BD057526	BD057526 Netrin re
32	735.8	25.5	813	6	AX054892	AX054892 Sequence
33	665.6	23.1	2832	5	AY099459	AY099459 Xenopus 1
34	597.6	20.7	2661	6	AX800717	AX800717 Sequence
35	597.6	20.7	2668	6	AX800719	AX800719 Sequence
36	573.4	19.9	604	6	AX451656	AX451656 Sequence
37	560.6	19.5	3252	9	AY358147	AY358147 Homo sapi
38	557.8	19.4	7002	6	BD162193	BD162193 Novel unc
39	557.8	19.4	7002	9	AB055056	AB055056 Homo sapi
40	553	19.2	3925	10	MMU487854	MMU487854 Mus muscu
41	502.6	17.4	1438	6	AX054888	AX054888 Sequence
42	467.4	16.2	2406	6	AX833115	AX833115 Sequence
43	467.4	16.2	2406	9	AK094595	AK094595 Homo sapi
44	465.8	16.2	1948	6	AX882354	AX882354 Sequence
45	465.8	16.2	1948	6	BD159676	BD159676 Primer fo

ALIGNMENTS

RESULT 1
AX527916
LOCUS AX527916 2881 bp DNA linear PAT 21-NOV-2002
DEFINITION Sequence 1 from Patent WO0229038.
ACCESSION AX527916
VERSION AX527916.1 GI:25172359
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Herrmann, J. L., Rastelli, L. and Shimkets, R. A.
TITLE Novel proteins and nucleic acids encoding same and antibodies
directed against these proteins

Db 1981 TGGGCCGCTTGGCCCTGGTGGAGAGGCCCTCAGCGTGGCTGCGGCCAAGCGCTCAAGC 2040
Qy 2041 TGCCTTCTGTTTGGCCCGGTGGCTTGCACCTCCCTCGAGTCAACATCCGGTCTACTGCC 2100
Db 2041 TGCCTTCTGTTTGGCCCGGTGGCTTGCACCTCCCTCGAGTCAACATCCGGTCTACTGCC 2100
Qy 2101 TGCATGACACCCAGATGCACTCAAGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGAC 2160
Db 2101 TGCATGACACCCAGATGCACTCAAGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGAC 2160
Qy 2161 AGCTGATCCAGGAGCCAGGGTCTGCACCTTCAAGGACAGTTACCAAACTTGCCTTAT 2220
Db 2161 AGCTGATCCAGGAGCCAGGGTCTGCACCTTCAAGGACAGTTACCAAACTTGCCTTAT 2220
Qy 2221 CCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAAGCTCCCTGTGAGTACACAGAGA 2280
Db 2221 CCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAAGCTCCCTGTGAGTACACAGAGA 2280
Qy 2281 TCCCTTTTATCATCTGGAATGGACGACGCGGTACTTGCACCTGCACCTTACCCCTGG 2340
Db 2281 TCCCTTTTATCATCTGGAATGGACGACGCGGTACTTGCACCTGCACCTTACCCCTGG 2340
Qy 2341 AGCGTGTACGCCAGCAGTACTAGTACCTGGCTGCAAGCTGTGGTGTGCAAGTGGAGG 2400
Db 2341 AGCGTGTACGCCAGCAGTACTAGTACCTGGCTGCAAGCTGTGGTGTGCAAGTGGAGG 2400
Qy 2401 GCGACGGGACAGAGCTTCAAGCATCAACTTCAACATCACCAAGGACACAAAGTTTGTGAGC 2460
Db 2401 GCGACGGGACAGCTTCAAGCATCAACTTCAACATCACCAAGGACACAAAGTTTGTGAGC 2460
Qy 2461 TGTGTGCTCTGGAGAGTGAAGCGGGGTCCAGCCCTGTGGTGGCCCTTCAAGA 2520
Db 2461 TGTGTGCTCTGGAGAGTGAAGCGGGGTCCAGCCCTGTGGTGGCCCTTCAAGA 2520
Qy 2521 TCCCTTCTCATTCGGCAGAGATATTTCCAGCTGGACCCACCTGTAGCGGGGTG 2580
Db 2521 TCCCTTCTCATTCGGCAGAGATATTTCCAGCTGGACCCACCTGTAGCGGGGTG 2580
Qy 2581 CCGACTGGCGGACTCTGGGCCAGAACTCCACCTGGACAGCCATCTCAGCTTCTTGGCT 2640
Db 2581 CCGACTGGCGGACTCTGGGCCAGAACTCCACCTGGACAGCCATCTCAGCTTCTTGGCT 2640
Qy 2641 CCAAGCCAGCCCAAGCCATGATCCTCAACTGTGGAGGCGCGCATCTCCCCCAAG 2700
Db 2641 CCAAGCCAGCCCAAGCCATGATCCTCAACTGTGGAGGCGCGCATCTCCCCCAAG 2700
Qy 2701 GCAACCTCAGCAGCTGGCTGAGCAGTGGTGGACTGGGCGAGCCAGCGTGGCTCT 2760
Db 2701 GCAACCTCAGCAGCTGGCTGAGCAGTGGTGGACTGGGCGAGCCAGCGTGGCTCT 2760
Qy 2761 TCACAGTGTGGAGGCTGAGTGTGAGGCGCGCCAGCGCGCATCTACCTTACACAG 2820
Db 2761 TCACAGTGTGGAGGCTGAGTGTGAGGCGCGCCAGCGCGCATCTACCTTACACAG 2820
Qy 2821 CTTTGGCACCACCAAGGACAGCGAGAACCCGAGAGGGGCGCTTCCCCACACCGGGAG 2880
Db 2821 CTTTGGCACCACCAAGGACAGCGAGAACCCGAGAGGGGCGCTTCCCCACACCGGGAG 2880
Qy 2881 A 2881
Db 2881 A 2881

RESULT 2
LOCUS AX449572
DEFINITION Sequence 1 from Patent WO0210216.
ACCESSION AX449572
VERSION AX449572.1 GI:21698195
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE
AUTHORS Padigaru,M., Mezes,P., Mishra,V., Burgess,C., Casman,S.,
Grosse,W.M., Alsobrook,J.P., Lepley,D.M., Gerlach,V.L.,
Macdougall,J.R. and Smithson,G.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 0210216-A 1 07-FEB-2002;
Curagen Corporation (US)
FEATURES
source 1. 2752
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 92.9%; Score 2676.4; DB 6; Length 2752;
Best Local Similarity 98.9%; Pred.No. 0;
Matches 2728; Conservative 0; Mismatches 21; Indels 9; Gaps 3;
Qy 42 CCGCGGGGCCCCCGCGCCCGCCCGCTGCGCCGCGCCGCGCCATGGCGCTCCGCGCC 101
Db 1 CCGCGGGGCCCCCGCGCCCGCCCGCTGCGCCGCGCCGCGCCATGGCGCTCCGCGCC 60
Qy 102 GGCCTGTGGCCAGCGCTCTTGGGCATAGTCTCGCCGCTTGGCTTCGCGGCTCGGCTGCC 161
Db 61 GGCCTGTGGCCAGCGCTCTTGGGCATAGTCTCGCCGCTTGGCTTCGCGGCTCGGCTGCC 120
Qy 162 CAGCAGAGTGGCCACCGTGGCCAAACCCAGTGCCTGGTGGCCAAACCCGACCTGCTCCCCAC 221
Db 121 CAGCAGAGTGGCCACCGTGGCCAAACCCAGTGCCTGGTGGCCAAACCCGACCTGCTCCCCAC 180
Qy 222 TTCTCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAAACAAAGCCAGTGTCTTGTGTGC 281
Db 181 TTCTCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAAACAAAGCCAGTGTCTTGTGTGC 240
Qy 282 AAGGCGCTGCCCGCCAGCAGATCTTCTCAAGTGAACGGGGAGTGGGTGGCCAGGTG 341
Db 241 AAGGCGCTGCCCGCCAGCAGATCTTCTCAAGTGAACGGGGAGTGGGTGGCCAGGTG 300
Qy 342 GACCACGTGATCGAGCGCAGCAGACGAGCGAGTGTGAGCCGACCATGGAGGTCCGC 401
Db 301 GACCACGTGATCGAGCGCAGCAGACGAGCGAGTGTGAGCCGACCATGGAGGTCCGC 360
Qy 402 ATTAATGTCTAAGCGCAGCAGTTCGAGAAAGTGTTCGGGCTGGAGGAATACTGGTCCAG 461
Db 361 ATTAATGTCTAAGCGCAGCAGTTCGAGAAAGTGTTCGGGCTGGAGGAATACTGGTCCAG 420
Qy 462 TCGGTGGCATGAGCTCTCGGGCACCAACAGAGTCAAGAGCTCAAGGCTTACATCCGCATAGCC 521
Db 421 TCGGTGGCATGAGCTCTCGGGCACCAACAGAGTCAAGAGCTCAAGGCTTACATCCGCATAGCC 480
Qy 522 AGATTGGCAGAAATTCGAGCAGGAGCGCTGGCCAGGAGGTGTCCCTGGAGCAGGGC 581
Db 481 AGATTGGCAGAAATTCGAGCAGGAGCGCTGGCCAGGAGGTGTCCCTGGAGCAGGGC 540
Qy 582 ATCTGTCTGCCCTGCCCTCCACCGGAGGCGCATCCCTCCAGCGAGGTGGAGTGGTCCCG 641
Db 541 ATCTGTCTGCCCTGCCCTCCACCGGAGGCGCATCCCTCCAGCGAGGTGGAGTGGTCCCG 600
Qy 642 AACGAGGACCTGGTGGACCGCTTGTGACACGCCCAATGTATATACATCACGCGGAGCACAGC 701
Db 601 AACGAGGACCTGGTGGACCGCTTGTGACACGCCCAATGTATATACATCACGCGGAGCACAGC 660
Qy 702 TGTGTGGTGGCAGCGCCCGCTTGTGACACGCCCAATGTATATACATCACGCGGAGCACAGC 761
Db 661 TGTGTGGTGGCAGCGCCCGCTTGTGACACGCCCAATGTATATACATCACGCGGAGCACAGC 720
Qy 762 ATCTGGCAGCTCGCGCAGCGCTTCCGCTGTCTATCTGCTAGTGAACGGTGGGTGG 821
Db 721 ATCTGGCAGCTCGCGCAGCGCTTCCGCTGTCTATCTGCTAGTGAACGGTGGGTGG 780
Qy 822 TCGAGTGGACCGAGTGGTCTGTGACGCGCCAGCTGTGGGCGCGCTGGCAGAAACGG 881

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

JOURNAL Patent: WO 0233080-A 1 25-APR-2002;
 BAYER Aktiengesellschaft (DE)
 FEATURES Location/Qualifiers
 source 1..2697
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN

Query Match 90.8%; Score 2615; DB 6; Length 2697;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 2669; Conservative 0; Mismatches 25; Indels 9; Gaps 3;
 87 ATGGCCGCTCCGCGCGCGCTGTGGCCAGCGCTCTCTGGGCAAGTCTCTGGCGCTTGGCTC 146
 Db 1 ATGGCCGCTCCGCGCGCGCTGTGGCCAGCGCTCTCTGGGCAAGTCTCTGGCGCTTGGCTC 60
 147 GCGCGCTCGGCTGCGCCAGCAGAGTGCACCGTGTGCCAACCCAGTGCCTGTGTGCCAACCCG 206
 Db 61 GCGCGCTCGGCTGCGCCAGCAGAGTGCACCGTGTGCCAACCCAGTGCCTGTGTGCCAACCCG 120
 207 GACCTGTCTCCCACTTCTCTGAGCGCGGAGTGTACATCGTCAAGAACCAAGCCA 266
 Db 121 GACCTGTCTCCCACTTCTCTGAGCGCGGAGTGTACATCGTCAAGAACCAAGCCA 180
 267 GTGCTGTCTGTGTGCAAGGCGCTGCCCGCCACGAGATCTTCTTCAAGTGCACCGGGAG 326
 Db 181 GTGCTGTCTGTGTGCAAGGCGCTGCCCGCCACGAGATCTTCTTCAAGTGCACCGGGAG 240
 327 TGGGTGCGCCAGGTGTGACACGATCGAGCGCAGCAGACGAGCGGAGCAGTGTGTGAGCCG 386
 Db 241 TGGGTGCGCCAGGTGTGACACGATCGAGCGCAGCAGACGAGCGGAGCAGTGTGTGAGCCG 300
 387 ACCATGAGGTGCGCATTAATGTCTCAAGCAGCAGGTTCAGAGAGTTCGGGCTGGAG 446
 Db 301 ACCATGAGGTGCGCATTAATGTCTCAAGCAGCAGGTTCAGAGAGTTCGGGCTGGAG 360
 447 GAA TACTGTGTGCCAGTGTGGATGAGCTCTCGGGCCACCAAGAGTTCAGAGGCC 506
 Db 361 GAATACTGTGTGCCAGTGTGGATGAGCTCTCGGGCCACCAAGAGTTCAGAGGCC 420
 507 TACATCGCATAGCAGATTGCGCAAGAACTTCGAGCAGGAGCGCTGGCCAAAGGAGTG 566
 Db 421 TACATCGCATAGCAGATTGCGCAAGAACTTCGAGCAGGAGCGCTGGCCAAAGGAGTG 480
 567 TCCTTGAGCAGGCGATCGTGTGCGCTGCGCTCCAGCGAGGCGATCCCTCCAGCGCGAG 626
 Db 481 TCCTTGAGCAGGCGATCGTGTGCGCTGCGCTCCAGCGAGGCGATCCCTCCAGCGCGAG 540
 627 GTGAGTGTGCTCGGAACGAGGACCTGTGTGGAACCGCTCTCGACCCCAATGTATATC 686
 Db 541 GTGAGTGTGCTCGGAACGAGGACCTGTGTGGAACCGCTCTCGACCCCAATGTATATC 600
 687 ACAGCGGAGCAGCGCTGTGTGTGAGAGCGCGCTGTGTGACACGCGCAACTACAC 746
 Db 601 ACAGCGGAGCAGCGCTGTGTGTGAGAGCGCGCTGTGTGACACGCGCAACTACAC 660
 747 TGGTGTGCGCAAGAACTTCGTGTGCAAGTTCGCGCAGCGCTCTCGCTGTATCTGTCTAC 806
 Db 661 TGGTGTGCGCAAGAACTTCGTGTGCAAGTTCGCGCAGCGCTCTCGCTGTATCTGTCTAC 720
 807 GTGAACGCTGGGTGTGTGAGTGTGGAACCGAGTGTGTGTGTGAGCGCCAGCTGTGGCGC 866
 Db 721 GTGAACGCTGGGTGTGTGAGTGTGGAACCGAGTGTGTGTGTGAGCGCCAGCTGTGGCGC 780
 867 GGTGTGCAAGAACGAGCGGAGTGTGCAACACCGCGCGCTCTCAACGCGGCGCTTTC 926
 Db 781 GGTGTGCAAGAACGAGCGGAGTGTGCAACACCGCGCGCTCTCAACGCGGCGCTTTC 840
 927 TGTGAGGCGCAGAACTGTCTTACCGCAGCGTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGT 986
 Db 841 TGTGAGGCGCAGAACTGTCTTACCGCAGCGTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGT 997

QY 987 AGCTGGAGCCCGTGGAGCAAGTGTGTGGCTGTGGCTGTGAGCTGACACCACTGGCGGAGC 1046
 Db 988 AGCTGGAGCCCGTGGAGCAAGTGTGTGGCTGTGGCTGTGAGCTGACACCACTGGCGGAGC 957
 1047 CGTGAGTGTCTGTGACCCAGCACCAGCAACGAGGAGGAGTGTGAGGAGGAGTGTGAGGAGC 1106
 Db 958 CGTGAGTGTCTGTGACCCAGCACCAGCAACGAGGAGGAGTGTGAGGAGGAGTGTGAGGAGC 1017
 1107 GACACCCGCAACTGTATACAGTGTGCTGTGTACACAGTGTCTGTGGCTCTGAGGAGCTG 1166
 Db 1018 GACACCCGCAACTGTATACAGTGTGCTGTGTACACAGTGTCTGTGGCTCTGAGGAGCTG 1077
 1167 GCGCTCTATGTGGCTCTATGCGCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1226
 Db 1078 GCGCTCTATGTGGCTCTATGCGCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1137
 1227 ATCTCTGTTTATTTGCGGAAAGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1286
 Db 1138 ATCTCTGTTTATTTGCGGAAAGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1197
 1287 CTCACTCAGGCTTCCAGCCCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1346
 Db 1198 CTCACTCAGGCTTCCAGCCCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1257
 1347 CTCACTCAGGCTTCCAGCCCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1403
 Db 1258 CTCACTCAGGCTTCCAGCCCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1317
 1404 CGGAGGATGGGCGCCAGCCCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1463
 Db 1318 CGGAGGATGGGCGCCAGCCCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1377
 1464 GGTGGGCGCCGACACACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1523
 Db 1378 GGTGGGCGCCGACACACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1437
 1524 TCCGCGCTTCCAGCCAGAACTTCTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGT 1583
 Db 1438 TCCGCGCTTCCAGCCAGAACTTCTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGT 1497
 1584 TATGGAGCTTCAACTTCTCGGGGCGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1643
 Db 1498 TATGGAGCTTCAACTTCTCGGGGCGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1557
 1644 CTCACTCCCGCAGATGCCATACCCCGAGGAGAACTTATGAGATCTACCTCACTGAGTGTGAGT 1703
 Db 1558 CTCACTCCCGCAGATGCCATACCCCGAGGAGAACTTATGAGATCTACCTCACTGAGTGTGAGT 1617
 1704 NAGCGGAGAGCGTGTGAGTGTGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1763
 Db 1618 NAGCGGAGAGCGTGTGAGTGTGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1677
 1764 AGCTGTGGAGCCCGCTGGCGTCTGTGCTCACCGGCGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1823
 Db 1678 AGCTGTGGAGCCCGCTGGCGTCTGTGCTCACCGGCGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1737
 1824 GGGAGCCAGCCCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1883
 Db 1738 GGGAGCCAGCCCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1797
 1884 TGGGAGCAGGATGTGTGCACTGTGGCGGAGGAGGCGCTTCCCACTCTACTACTACTACTACT 1943
 Db 1798 TGGG---AGGATGTGTGCACTGTGGCGGAGGAGGCGCTTCCCACTCTACTACTACTACTACT 1854
 1944 CTGGAGGCGCAGT 2003
 Db 1855 CTGGAGGCGCAGT 1914
 2004 GAGGCGCTCAGCGT 2063
 Db 1915 GAGGCGCTCAGCGT 1974
 2064 TGCACTCTCTCGAGTGTGCACTCGGGGTCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2123

Db	1975	TGCACTCCCTCGATTACAACTCGGGTCTACTGCTGCATGACACCCACGATGCACCTC	2034
Qy	2124	AAGAGGTGTGTGAGCTGGAGAAAGCAGCTGGGGGGAAGCTGATCCAGAGACCACGGGTTC	2183
Db	2035	AAGGAGTGTGTGAGCTGGAGAAAGCAGCTGGGGGGAAGCTGATCCAGAGACCACGGGTTC	2094
Qy	2184	CTGCATTCAAGAGACAGTTACCAACCTTGGCTATCCATCCACGATGTGCCAGCTCC	2243
Db	2095	CTGCATTCAAGAGACAGTTACCAACCTTGGCTATCCATCCACGATGTGCCAGCTCC	2154
Qy	2244	CTGTGGAAGAGTAAAGTCCTTTGTTCAGCTACAGAGATCCCTTTTATCACATCTGGAAT	2303
Db	2155	CTGTGGAAGAGTAAAGTCCTTTGTTCAGCTACAGAGATCCCTTTTATCACATCTGGAAT	2214
Qy	2304	GGCAGCAGCGGTACTTTGCATCTGCACCTTCAGAGCGTGTACGCCCCAGCATAGT	2363
Db	2215	GGCAGCAGCGGTACTTTGCATCTGCACCTTCAGAGCGTGTACGCCCCAGCATAGT	2274
Qy	2364	GACTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGGCAACGGGAGAGCTTCAGCATC	2423
Db	2275	GACTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGGCAACGGGAGAGCTTCAGCATC	2334
Qy	2424	AACTTCAACATCACCAGAGCACAAAGCTTTGCTGAGCTGTGGCTCTGGAGAGTGAAGCG	2483
Db	2335	AACTTCAACATCACCAGAGCACAAAGCTTTGCTGAGCTGTGGCTCTGGAGAGTGAAGCG	2394
Qy	2484	GGGTGCCAGCCCTGTGTGGGCCCCAGTGCCTTCAAAGATCCCTTCCTCATTCGGGCAAG	2543
Db	2395	GGGTGCCAGCCCTGTGTGGGCCCCAGTGCCTTCAAAGATCCCTTCCTCATTCGGGCAAG	2454
Qy	2544	ATAATTTCCAGCCTGGAACCAACCTGTAGGGGGGTGCCGAGTGGCGGAATCTGGGCCAG	2603
Db	2455	ATAATTTCCAGCCTGGAACCAACCTGTAGGGGGGTGCCGAGTGGCGGAATCTGGGCCAG	2514
Qy	2604	AAACTCCACTGGAGACGCATCTCAGCTTCTTTGCTTCCAAAGCCCGCCACAGCCATG	2663
Db	2515	AAACTCCACTGGAGACGCATCTCAGCTTCTTTGCTTCCAAAGCCCGCCACAGCCATG	2574
Qy	2664	ATCCTCAACCTGTGGGAGGGCGGCACCTTCCCAACGGCAACCTCAGCAGCTGGCTGCA	2723
Db	2575	ATCCTCAACCTGTGGGAGGGCGGCACCTTCCCAACGGCAACCTCAGCAGCTGGCTGCA	2634
Qy	2724	GCAGTGGCTGACCTGGGCGACGAGACGCTGGCTCTTTCACAGTGTGGAGGCTGAGTGC	2783
Db	2635	GCAGTGGCTGACCTGGGCGACGAGACGCTGGCTCTTTCACAGTGTGGAGGCTGAGTGC	2694
Qy	2784	TGA 2786	
Db	2695	TGA 2697	

[illegible]

```

source
1. .3580
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 6052371CB1"

ORIGIN
Query Match      83.4%; Score 2402.6; DB 6; Length 3580;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 2615; Conservative 0; Mismatches 9; Indels 177; Gaps 3;

QY 84 GCGATGCGCGTCCGGCCCGCCGCTGTGGCAGCGCTCTCTGGGCATAGTCTCTCGCGCGCTTGG 143
DB 1 GCCATGCGCGTCCGGCCCGCCGCTGTGGCAGCGCTCTCTGGGCATAGTCTCTCGCGCGCTTGG 60

QY 144 CTCGCGCGCTCGGCTGCCAGCAGAGTGCACCGTGGCCCAACCCAGTGCCTGTGTGGCCCAAC 203
DB 61 CTCGCGCGCTCGGCTGCCAGCAGAGTGCACCGTGGCCCAACCCAGTGCCTGTGTGGCCCAAC 120

QY 204 CCGGACCTGCTTCCCACTTCTCTGGTGGAGCCGAGGATGTGTACATCTGTCAGAGCAACAG 263
DB 121 CCGGACCTGCTTCCCACTTCTCTGGTGGAGCCGAGGATGTGTACATCTGTCAGAGCAACAG 180

QY 264 CCAGTGTCTTGTGTGTCAGAGCGCTGCCCGCACCGCAGATCTTCTCAAGTGCACAGGG 323
DB 181 CCAGTGTCTTGTGTGTCAGAGCGCTGCCCGCACCGCAGATCTTCTCAAGTGCACAGGG 240

QY 324 GAGTGGGTGCGCCAGGTGGACCGTCTATCGAGCGCAGCAGACGCGGACGAGTGTGTGAG 383
DB 241 GAGTGGGTGCGCCAGGTGGAGCACCGTCTATCGAGCGCAGCAGTGTGTGAGTGTGTGAG 300

QY 384 CCGACCATGAGGTCGCGCATTTAATGTCTCAAGGCAGCAGTTCGAGAGGTGTGTGGGCTG 443
DB 301 CCCACCATGAGGTCGCGCATTTAATGTCTCAAGGCAGCAGTTCGAGAGGTGTGTGGGCTG 360

QY 444 GAGGAATACTGGTGCAGTGGCTGGCATGAGCTCTCTCGGGCACCACCAAGAGTTCAGAAG 503
DB 361 GAGGAATACTGGTGCAGTGGCTGGCATGAGGCTCTCTCGGGCACCACCAAGAGTTCAGAAG 420

QY 504 GCCTACATCCGCATAGCAGATTCCGCAAGAATCTCGAGCAGGAGCGCTGGCCCAAGGAG 563
DB 421 GCCTACATCCGCATAGCCTATTTCGCAAGAACTTCGAGCAGGAGCGCTGGCCCAAGGAG 480

QY 564 GTGTCCTCTGGAGCAGGGGCATCGTGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCC 623
DB 481 GTGTCCTCTGGAGCAGGGGCATCGTGTGCTGCCCTGCCGTCCACCGGAGGGGCATCCCTCCAGCC 540

QY 624 GAGGTGGAGTGGCTCCGGAAACGAGAACCTGTGTGACCCGCTCCCTGGACCCCAATGTATAC 683
DB 541 GAGGTGGAGTGGCTCCGGAAACGAGAACCTGTGTGACCCGCTCCCTGGACCCCAATGTATAC 600

QY 684 ATCAGCGGGAGCAGACCTGCTGTGCGACAGGCGCGCTTGTGTGACACGGGCCAACTAC 743
DB 601 ATCAGCGGGAGCAGACCTGCTGTGCGACAGGCGCGCTTGTGTGACACGGGCCAACTAC 660

QY 744 ACCTCGCTGGCCAAAGAACATCGTGGCAGTGCCTCGCAGCGCTCCGCTGTGTCTCATGTC 803
DB 661 ACCTCGCTGGCCAAAGAACATCGTGGCAGTGCCTCGCAGCGCTCCGCTGTGTCTCATGTC 720

QY 804 TACGTGAACCGTGGGTGTGACAGTGTGACCGAGTGTCCGCTCTGACGGGCCAGCTGTGGG 863
DB 721 TAC-----TAC-----TAC-----TAC-----TAC-----TAC----- 723

QY 864 CCGGCTGGCAGAACGGAGCCGGAGTGCACCAACCCGGCGCTCTCAACGGGGGGCGCT 923
DB 724 -----TAC-----TAC-----TAC-----TAC-----TAC-----TAC----- 723

QY 924 TTCTGTGAGGGCAGAAATGTCCATGACCGCACCGCTCTCTCTGTCTTGTCTCTGTGAC 983
DB 724 -----TAC-----TAC-----TAC-----TAC-----TAC-----TAC----- 729

QY 984 GGCAGCTGAGCCCTGTGAGCAAGTGTGTGGCCTGTGGGCTGGAGCTGCACCCACTTGGCGG 1043
DB 724 -----TAC-----TAC-----TAC-----TAC-----TAC-----TAC----- 1043

```

Db 730 GGAGCTGGAGCCCGTGGAGCAAGTGTGCGCTGTGGCTGGACTGACACCACTGGCGG 789
QY 1044 AGCCGTGAGTGTCTCTGACCCAGCACCCCGCAACGAGGAGGAGTGCAGGGCACTGAC 1103
Db 790 AGCCGTGAGTGTCTCTGACCCAGCACCCCGCAACGAGGAGGAGTGCAGGGCACTGAC 849
QY 1104 CTGAGACCCCGCAACTGTPACAGTGAACCTCTGTGTACACAGTGTCTTGGCCCTGAGGAC 1163
Db 850 CTGAGACCCCGCAACTGTPACAGTGAACCTCTGTGTACACAGTGTCTTGGCCCTGAGGAC 909
QY 1164 GTGGCCCTCTATGTGGGCTCATCGCCGTGGCCGTCTGCTGGTCTCTGCTGCTGTGTC 1223
Db 910 GTGGCCCTCTATGTGGGCTCATCGCCGTGGCCGTCTGCTGGTCTCTGCTGCTGTGTC 969
QY 1224 CTCATCCTCGTTATTGCGGGAAGAAGAGGGGTGGACTCAGATGTGGCTGACTCGTCC 1283
Db 970 CTCATCCTCGTTATTGCGGGAAGAAGAGGGGTGGACTCAGATGTGGCTGACTCGTCC 1029
QY 1284 ATTCTACCTCAGGCTTCCAGCCCGTCAAGCTCAGCATCAAGGCCAGCAAGCAACCCCAT 1343
Db 1030 ATTCTACCTCAGGCTTCCAGCCCGTCAAGCTCAGCATCAAGGCCAGCAAGCAACCCCAT 1089
QY 1344 CTGCTACCATCCAGCCGAGCTCAG--CACACACACAGCTTACAGGGAGTCTGT 1400
Db 1090 CTGCTACCATCCAGCCGAGCTCAGCACCACCAACCACTACAGGGAGTCTGT 1149
QY 1401 CCCGGCAGGATGGGCCCGAGCCCAAGTTCAGCTCACAATGGGACCTGTCTCAGCCCC 1460
Db 1150 CCCGGCAGGATGGGCCCGAGCCCAAGTTCAGCTCACAATGGGACCTGTCTCAGCCCC 1209
QY 1461 CTGGTGGCGGCGGCGACACACTGCAACACAGTCTCCACCTCTGAGGCGGAGGATTC 1520
Db 1210 CTGGTGGCGGCGGCGACACACTGCAACACAGTCTCCACCTCTGAGGCGGAGGATTC 1269
QY 1521 GTCTCCGCTCTCCACCCAGACTACTTCGGTCCCTGCGCCGAGGACACAGCAACG 1580
Db 1270 GTCTCCGCTCTCCACCCAGACTACTTCGGTCCCTGCGCCGAGGACACAGCAACG 1329
QY 1581 ACCTATGGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTATACAGTATCAGC 1640
Db 1330 ACCTATGGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTATACAGTATCAGC 1389
QY 1641 CTCTCATCCCCCAGATGCCATACCCCGAGGAAGATCTATGAGATCTACCTCAGCTG 1700
Db 1390 CTCTCATCCCCCAGATGCCATACCCCGAGGAAGATCTATGAGATCTACCTCAGCTG 1449
QY 1701 CACAAGCGGAGAGCTGAGTTCGCCCTAGCTGGCTGTGACACCTGTGAGTCCCATC 1760
Db 1450 CACAAGCGGAGAGCTGAGTTCGCCCTAGCTGGCTGTGACACCTGTGAGTCCCATC 1509
QY 1761 GTTAGTGTGGACCCCTCGCGTCTCTGCTCAGCCGCGAGTATCTCTGGCTATGGACCAC 1820
Db 1510 GTTAGTGTGGACCCCTCGCGTCTCTGCTCAGCCGCGAGTATCTCTGGCTATGGACCAC 1569
QY 1821 TGTGGGAGCCAGCCCTGACAGCTGAGCCTCGGCTCAAAAAGCAGTGTGCGAGGGC 1880
Db 1570 TGTGGGAGCCAGCCCTGACAGCTGAGCCTCGGCTCAAAAAGCAGTGTGCGAGGGC 1629
QY 1881 AGCTGGAGCAGATGTGCTGACCTGGGAGGAGGCGCCCTCCACCTCTACTACTG 1940
Db 1630 AGCTGGG--AGAGTGTGTGACCTGGGAGGAGGCGCCCTCCACCTCTACTACTG 1686
QY 1941 CAGCTGGAGCCAGTGCCTGTACGTCTTACCGAGCAGTGGCGCTTCCCTGCTG 2000
Db 1687 CAGCTGGAGCCAGTGCCTGTACGTCTTACCGAGCAGTGGCGCTTCCCTGCTG 1746
QY 2001 GGAGAGCCCTCAGCGTGTGCGCCAGCGCTCAAGCTGTCTGTTTGGCGCGTG 2060
Db 1747 GGAGAGCCCTCAGCGTGTGCGCCAGCGCTCAAGCTGTCTGTTTGGCGCGTG 1806
QY 2061 GCTGTCACTCCCTCAGTATCAACATCCGGGTCTACTGCTGATGACACCAACGATGA 2120
Db 1807 GCTGTCACTCCCTCAGTATCAACATCCGGGTCTACTGCTGATGACACCAACGATGA 1866

QY 2121 CTCAGAGAGTGTGTGAGTGTGAGAAAGCAGCTGGGGGACAGCTGATCCAGAGCCACGG 2180
Db 1867 CTCAGAGAGTGTGTGAGTGTGAGAAAGCAGCTGGGGGACAGCTGATCCAGAGCCACGG 1926
QY 2181 GTCTGTGACATTCAGGACAGTTCACCAACCTGCGCTTATCCATCCAGATGTGCCACG 2240
Db 1927 GTCTGTGACATTCAGGACAGTTCACCAACCTGCGCTTATCCATCCAGATGTGCCACG 1986
QY 2241 TCCCTGTGGAGAGTAAGTCTCTTGTGAGTCAACAGAGATCCCTTTTATCATCTGG 2300
Db 1987 TCCCTGTGGAGAGTAAGTCTCTTGTGAGTCAACAGAGATCCCTTTTATCATCTGG 2046
QY 2301 AATGGCAGCAGCGGTACTTGCACTGCACTTCAACCTGGAGCGTGTGAGCGGAGCTT 2360
Db 2047 AATGGCAGCAGCGGTACTTGCACTGCACTTCAACCTGGAGCGTGTGAGCGGAGCTT 2106
QY 2361 AGTGACCTGCGCTGCAAGCTGTGGCTGTGAGTGTGAGCGGAGCGGAGCGGAGCTT 2420
Db 2107 AGTGACCTGCGCTGCAAGCTGTGGCTGTGAGTGTGAGCGGAGCGGAGCGGAGCTT 2166
QY 2421 ATCAACTTCAACATCACCAAGGACACAAGTGTGCTGAGCTGTGGCTCTGGAGAGTGA 2480
Db 2167 ATCAACTTCAACATCACCAAGGACACAAGTGTGCTGAGCTGTGGCTCTGGAGAGTGA 2226
QY 2481 GCGGGGTCCAGCCCTGTGGTGGGCGCCAGTGTCTTCAAGATCCCTTCTCATTCGGCAG 2540
Db 2227 GCGGGGTCCAGCCCTGTGGTGGGCGCCAGTGTCTTCAAGATCCCTTCTCATTCGGCAG 2286
QY 2541 AAGATTAATTTCCAGCTTGGACCCACCTGTGAGCGGGGTGCGGAGCTTGGGCTTGGCC 2600
Db 2287 AAGATTAATTTCCAGCTTGGACCCACCTGTGAGCGGGGTGCGGAGCTTGGGCTTGGCC 2346
QY 2601 CAGAAATCCACCTTGACAGCCATCTCAGCTTCTTTGCTTCCAAAGCCAGCCGCCACAGCC 2660
Db 2347 CAGAAATCCACCTTGACAGCCATCTCAGCTTCTTTGCTTCCAAAGCCAGCCGCCACAGCC 2406
QY 2661 ATGATCCTTAACTGTGGAGGCGGCGACTTCCCAAGCCAGCCGCCACAGCCGCCCT 2720
Db 2407 ATGATCCTTAACTGTGGAGGCGGCGACTTCCCAAGCCAGCCGCCACAGCCGCCCT 2466
QY 2721 GCAGCAGTGTGAGTGTGGCCAGCCAGCGCTTGTGCTTCAAGTGTGCGAGGCTGAG 2780
Db 2467 GCAGCAGTGTGAGTGTGGCCAGCCAGCGCTTGTGCTTCAAGTGTGCGAGGCTGAG 2526
QY 2781 TGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2840
Db 2527 TGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2586
QY 2841 AGGCAGAACCGGACAGGGGCGGCTTCCCAAGCCAGCGGGGAGA 2881
Db 2587 AGGCAGAACCGGACAGGGGCGGCTTCCCAAGCCAGCGGGGAGA 2627

RESULT 5

MMU487852

LOCUS

DEFINITION

AJ487852

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

MMU487852 3992 bp mRNA linear ROD 24-SEP-2002
Mus musculus mRNA for netrin receptor Unc5h1 (Unc5h1 gene).

AJ487852.1 GI:22035783
netrin receptor Unc5h1; Unc5h1 gene.

Mus musculus
Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Engelkamp,D.

Cloning of three mouse Unc5 genes and their expression patterns at
mid-gestation

Mech. Dev. 118 (1-2), 191-197 (2002)

12351186

2 (bases 1 to 3992)

AUTHORS Engelkamp, D.
 TITLE Direct Submission
 JOURNAL Submitted (15-MAY-2002) Neuroanatomy, Max Planck Institute for Brain Research, Deutscherordenstrasse 46, Frankfurt 60528, GERMANY

FEATURES
 source /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 1. .3992
 /gene="Unc5h1"
 232..2928
 /gene="Unc5h1"
 /product="netrin receptor Unc5h1"
 /protein_id="CAD32250.1"
 /db_xref="GI:22035784"
 /db_xref="GOA:O8K154"
 /db_xref="SPTREMBL:O8K154"
 /translation="MAVRPLGLWALLIGVLTWLRGSAQOQSATVANPVPANPDLLP
 HFLVPEDEVIVKKNPVLLVCAVPATQIFKCNGEWVRQDVHIERSTDGSSGLPTM
 EVRINVSQVQVFLGLEEYWCQVAMSSGTTKSKQAYIRIAYLRNFFQEPLEKEV
 SLEQGVLPQCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQALADTAN
 YTCVAKNIVARRSASAAVIVVNGWSWTWESVCSACGRQWKRSCRITNPAPLN
 GGAPCGQNVQKTACATLCPVDGSMSPWSKAGLDCDTHWRSRQCSDPAPRNGBEC
 RGADLDRNCTSLCLHTSSGPEDVALYIGLVAVAVCLILLLLVLLVLYCRKEGLDS
 DVADSLTSGFQVPSIKPSKADNPILLITQDLSITTTTQSLCPRGQPSKPLQL
 SNGLHSLPLSGRHTLHSSPTSEAEFVSRSLSTONYFRSLPRGTSMAYGTENFELGG
 RLMPNTGISLLIPDPAIPRKIVIEYILTHKPEDVRLPLAGCQTLPLSPVSCGPQV
 LLTRPVTIAMDHCEGPSPDSRLRKQSCGSEWEDVHLGERSPSHLYYCOLEAGAC
 YVTEQLGRPALVGEALSVAATKRLRLLPAPVACTSLVNIYRVLCHDTHDALKEV
 QLEKQLGQILQPRVLFKDSYHNRLSLHDVPSLSLWKSLLVSVQIEIPFYHWNKT
 QYVLTHTFTLERNVNASDLACKVWVQVDEGDSFNINFTNDTRFAEMLALESBG
 GVPLVGPSPAFKIPFLRQKIITSLDPPCSRGADWRTLAQKJLHDLHSLSPFASKPSPT
 AMTLNLWEARHFPNGNLQGLAAAVALGLQPDAGLFTVSEAE"

ORIGIN

Query Match 80.8%; Score 2327.8; DB 10; Length 3992;
 Best Local Similarity 89.3%; Pred. No. 0;
 Matches 2579; Conservative 0; Mismatches 292; Indels 16; Gaps 6;

QY 1 AGCTGGGCGCTCCGGCTGAGCGGT-AAAGCGCGCTCCCGCGCGGGGCGCGCGCC 59
 DB 149 AGGTGGGCTCTCGGTGAGCGGTGACAGCTCTCTCCCGCGCGGGGCGCGCTCC 208
 QY 60 GGGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 119
 DB 209 AGCCCGCTCGT-----CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 264
 QY 120 CTGGGCGATGCTCCCGCGCGGTGCTCCCGCGCGGTGCGCGCGCGCGCGCGCGCG 179
 DB 265 CTGGGCGATGCTCCTCACTCGCTGCGTCTCGTGGTGGCGCGCGCGCGCGCGCG 324
 QY 180 GCCAACCGAGTGGTGGCGCAACCGCGCGCGCTTCCCGCGCGCGCGCGCGCGCG 239
 DB 325 GCCAACCGAGTGGTGGCGCAACCGCGCGCGCTTCCCGCGCGCGCGCGCGCGCG 384
 QY 240 GATGTGATCATCGTCAAGAAACAGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 299
 DB 385 GACGTGTACATTGTCAAGAAACAGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGT 444
 QY 300 CAGATCTTCTCAAGTGCAAGGGGAGTGGGTGGCGCGCGCGCGCGCGCGCGCGCG 359
 DB 445 CAGATCTTCTCAAGTGCAAGGGGAGTGGGTGGCGCGCGCGCGCGCGCGCGCGCG 504
 QY 360 AGCAGACGGGAGCGAGTGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 419
 DB 505 AGCATGTGAGCGAGCGAGCGGATTGCCAACCATTGGAGGTCCGGATCAACGTAAC 564
 QY 420 CAGGTGAGAGGTTCTCGGGCTGAGGAAATACCTGGTGGCGAGTGGCGATGGAGTCC 479
 DB 565 CAGGTGAGAGGTTCTTTGGGCTGGAGGAGTACTGGTCCAGGTGTGTGTGTGTGTGT 624

QY 480 TGGGGCACCAAGAGTCCAGAGGCTACATCCGATAGCAGATTGGCAAGAACTTC 539
 DB 625 TCAGGAAACACCAAAAGCCAGAGGCTACATCCGATTGGCTATTGGCAAGAACTTT 684
 QY 540 GAGCAGAGCGCTGGCCAAAGAGGTGTCTCTGAGCAGGCGATCGTGTGCTGCTCCGT 599
 DB 685 GAGCAGAGCGCTGGCCAAAGAGGTGTCTCTGAGCAGGCGATTTGTGTACCTTGTGCG 744
 QY 600 CCACCGAGGGGATCCCTCCAGCCGAGGTGGCTCCGGAACAGAGGACCTTGTGTGAC 659
 DB 745 CCCCCGGAAGGAATCCCCCAGCTGAGGTGGGTGGCTCCGAATAGAGGACCTCTGTG 804
 QY 660 CCGTCTCTGACCGCCCAATGTATACATCAGCGGAGCAGAGCCTGTGTGTGCGACAG 719
 DB 805 CCGTCTCTGACCGCCCAATGTGTATACATCAGCGGAGCAGAGCCTGTGTGTGCGAG 864
 QY 720 CGCTTGTGTGACCGGCGCAACTACCTGTGTGTGCGGAGCAGAGCCTGTGTGTGCGAC 899
 DB 865 CGCTTGTGTGACCGGCGCAACTACCTGTGTGTGCGGAGCAGAGCCTGTGTGTGCGAC 924
 QY 780 AGCGCTCTCGCTGTCTGTCTAGTGAAACGTTGGGTGGTTCGAGCTGGACCGAGTGG 839
 DB 925 AGCGCTCTCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 984
 QY 840 TCGCTCTGACCGCGCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 899
 DB 985 TCGCTCTGCACTGCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1044
 QY 900 CCGCGCGCTCTCAACGGGGCGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 959
 DB 1045 CCGCGCGCTCTCAACGGGGCGCGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1104
 QY 960 TCCTCTCTGT 1019
 DB 1105 GCCACTCTG---TGCCAGT 1161
 QY 1020 GGGCTGAGCTGCACCCACTGCGCGAGCGCTGAGTGTCTGTGACCGAGCAGCGCGCAAC 1079
 DB 1162 GGGCTTGTGCTGACCCCACTGCGCGAGCGCGGAGTGTCTCCGACCCAGCGCGCGCAAC 1221
 QY 1080 GGGGAGAGGTGCCAGGCGACTGACCTGTGACACCGCGAACTGTGTACAGTGTGTGTGT 1139
 DB 1222 GGTGAGGAGTGTGCGGGGT 1281
 QY 1140 CAGAGT 1199
 DB 1282 CACACCTTTTCGCGCGCGCGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1341
 QY 1200 TGGCTGT 1259
 DB 1342 TGGCTGT 1401
 QY 1260 GACTCAGATGT 1319
 DB 1402 GACTCAGAGT 1461
 QY 1320 CCCAGAGAGCAGAGCAACCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1376
 DB 1462 CCCAGAGAGCAGAGCAATCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1521
 QY 1377 ACCACCTACAGGCGAGT 1436
 DB 1522 ACCACCTACAGGCGAGCTGT 1581
 QY 1437 ACCATGGGACCTGT 1496
 DB 1582 TCTAATGGT 1641
 QY 1497 CCCACCTCTGAGGCGAGGAGT 1556
 DB 1642 CCCACCTCTGAGGCTGT 1701
 QY 1557 CTGCCCCGAGGACGACGACCAATGACCTATGGGACCTTCAACTTCTCGGGGCGCGGCTG 1616

Db 1702 CTGCCCCGGGTACAGCAACATGGCCTATGGACCTTCAACTTCTCGGGGGCGGCTG 1761
QY 1617 ATGATCCCTTAATACAGGTATCAGCTCTCTCATCCCCCAGATGCATACCCCGAGGAAG 1676
Db 1762 ATGATCCCTTAACACAGGAATCAGGCTCTCTCATACCCCGGACGCAATCCCGGAGGAAG 1821
QY 1677 ATCTATGAGATCTACCTCAGCTGCACAAGCGGAAGACGTGAGGTTGGCCCTAGCTGGC 1736
Db 1822 ATCTACGAGATCTACCTCACTCTGCACAGCAAGACGTGAGGTTGGCCCTAGCTGGC 1881
QY 1737 TGTGAGACCTGTGAGTCCCATCGTTAGTGTGGAACCCCTGTGCTGCTCTCACCCGG 1796
Db 1882 TGTGAGACCTGTGAGTCCCATCGTTAGTGTGGAACCCCTGTGCTGCTCTCACCCGG 1941
QY 1797 CCAGTATCTGCTGATGACCACTGTGGGAGGCCAGCCCTGACACTGAGGCTCGC 1856
Db 1942 CCAGTATCTGCTGATGACCACTGTGGGAGGCCAGCCCTGACACTGAGGCTCGC 2001
QY 1857 CTCAAAAGCAGTCTGTCGAGGGCAGCTGGGAGCAGATGTGCTGCACCTGGGCGAGGAG 1916
Db 2002 CTCAAAAGCAGTCTGTCGAGGGCAGCTGGG--AGGACGTGTGCACTTGTGAGGAG 2058
QY 1917 GCGCCCTCCACCTCTACTACTGCTCAGCTGGAGGCCAGTCCCTGCTACGCTTTCACCGAG 1976
Db 2059 TCGCCCTCTCATCTACTACTGCTCAGCTGGAGGCCAGTCCCTGCTATGCTTCAACCGAG 2118
QY 1977 CAGCTGGCGCTTTGCCCCCTGTGGGAGAGGCCCTCAGCGTGGTCCGCGCAGCGCTC 2036
Db 2119 CAGCTAGCGCTTTGCCCCCTGTGGGAGAGGCCCTCAGCGTGGTCCGCAACGAGCGCTC 2178
QY 2037 AAGCTCTTCTGTTGGCGCGTGGCTGCACCTCCCTCGAGTACACATCCCGGCTCTAC 2096
Db 2179 AGSCTCTTCTGTTGGCCCTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2238
QY 2097 TGCCTGATGACACCCAGTACCTCAAGAGAGTGTGAGTGGAGAGCAGTGGGG 2156
Db 2239 TGCCTGACGACACCCAGTACCTCAAGAGAGTGTGAGTGGAGAGCAGTGGGT 2298
QY 2157 GGACGCTGATCAGAGAGCCAGGGTCCCTGCACCTTCAAGACAGATTACACACCTTGGC 2216
Db 2299 GGACGCTGATCAGAGAGCCAGGGTCCCTGCACCTTCAAGACAGATTACACACCTTGGC 2358
QY 2217 CTATCCATCCAGATGTGCCAGCTCCCTGTGGAAGTAACTCTTGTGCTACCTACCG 2276
Db 2359 CTGTCCATCCAGAGCTGCCAGCTCCCTGTGGAAGCAAGTCTTGTGCTACCTACCG 2418
QY 2277 GAGATCCCTTTTATCATCTGGAATGGACGACGGGTACTTGCATCTGCACCTTACC 2336
Db 2419 GAGATCCCTTTTATCATCTGGAATGGACGACGGGTACTTGCATCTGCACCTTACC 2478
QY 2337 CTGGAGCGTGTAGCCCCAGCACTAGTACCTGGCTGCAAGCTGTGGTGTGGAGGTG 2396
Db 2479 CTGGAGCGGTAAATGCCAGACACGAGCTGCGCTGCAAGGTGTGGTGTGGAGGTG 2538
QY 2397 GAGGGGAGCGGAGAGCTTACGATCAACTTCAACATCAACAGGACACAAAGTTTGTCT 2456
Db 2539 GAGGGAGATGGACAGAGCTTCAACATCAACTTAACTCACTAAGGACACAGAGTTTGTCT 2598
QY 2457 GAGCTGTGCTCTGGAGAGTGAAGCGGGGTCCAGCCCTGTGGGCCCCAGTGCCTTC 2516
Db 2599 GAAATGCTGGCTCTGGAGAGTGAAGGGGGGTCCAGCCCTGTGGGCCCCAGTGCCTTC 2658
QY 2517 AAGATCCCTTCTCATCTGCGGAGAGATAATTTCCAGCTGGAACCCCTGTAGGCGG 2576
Db 2659 AAGATCCCTTCTCATCTGCGGAGAGATAATTTCCAGCTGGAACCCCTGTAGGCGG 2718
QY 2577 GGTGCGGAGCTGGCGGACTCTGGCCAGAACTCCACTGACAGCACTCTAGCTTCTTT 2636
Db 2719 GCGCGGAGCTGGGAACTCTAGCCAGAACTTCACTTGGAGAGCCATTTAGCTTCTTT 2778
QY 2637 GCCTCGAAGCCAGCCCAACAGCCATGATCTCAACTGTGGGAGCGCGGCACTTCCCC 2696

Db 2779 GCCTCCAGCCAGCCCTACAGCCATGATCTCAACCTATGGGAGGCGCGGCACTTCCCC 2838
QY 2697 AACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGCTGGCCAGCCAGAGCTGGC 2756
Db 2839 AACGGCAACCTCAGCCAGCTGGCCGAGCTGTGGCGGACTGGGCCAGCCAGATGCTGGC 2898
QY 2757 CTCTTACAGTGTGCGAGGCTGAGTGTGAGGCGCGGCGAGGCCCG--ACACCTACATCT 2814
Db 2899 CTCTTACAGTGTGCGAGGCTGAGTGTGAGGCGCGGCGAGGCCCGCTACGCTACATCT 2958
QY 2815 CACGAGTTTGGCACCCCAAGCAGGAGCGGAGCGGAGCGGCGCTTCCCCACAC 2874
Db 2959 CACGAGTTTGGCACCTGCGGAGGAGCGGAGCGGAGCGGCGCTTCCCCACAC 3018
QY 2875 GGGGAGA 2881
Db 3019 CCGGGA 3025

RESULT 6
BD057524
LOCUS Netrin receptors. 3014 bp DNA linear PAT 27-AUG-2002
DEFINITION BD057524
ACCESSION BD057524
VERSION BD057524.1 GI:22603130
KEYWORDS JP 2001505062-A/1.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 3014)
AUTHORS Lavigne,M.T., Leonardo,D.E., Hinck,L., Masu,M. and Masu,K.K.
TITLE Netrin receptors
JOURNAL Patent: JP 2001505062-A 1 17-APR-2001;
COMMENT THE REGENTS OF THE UNIV OF CALIFORNIA
PN JP 2001505062-A/1
PD 17-APR-2001
PF 19-FEB-1998 JP 1998536840
PR 19-FEB-1997 US 08/808982
PI MARC TESSIER LAVIGNE,DAVID E LEONARDO,LINDSAY HINCK,MASAYUKI
PI MASU,
PI KAZUKO KEINO MASU
PC C07K1/00,C07K14/00,C07K17/00,C07H21/02,C07H21/04,G01N33/53 CC
Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source
1..3014
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 77.6%; Score 2235.8; DB 6; Length 3014;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;
QY 87 ATGGCGCTCGGCGCGGCTGTGGCCAGCGCTTCCTGGGCATAGTCTCGCCGCTTGGCTC 146
Db 1 ATGGCGCTCGGCGCGGCTGTGGCCAGTCTCTCTGGGCATAGTCTCTCGCCGCTTGGCTT 60
QY 147 GCGGCTCGGCTCGGCGCGGCTGTGGCCAGGTCGCCAGCTGGCCCAACCCAGTGCCTGGTGCACCCG 206
Db 61 CGTGTTCGGGTGCGGCGCGGCTGTGGCCAGGTCGCCAGTGCCTAGTGCCTGGTGCACCC 120
QY 207 GACCTGCTTCCCCTTCTCTGCTGAGCGCCGAGGATGTACATCGTCAAGACAGCA 266
Db 121 GACCTGCTGCCCCCTTCTCTGCTAGAGCTGAGGACGCTGATCATTTCAAGAACAGCG 180
QY 267 GTGCTGCTGTGTGCAAGCGCGTGCCTCGCCAGCAGATCTTCTTCAAGTGAACCGGGAG 326
Db 181 GTGTTGTTGGTGTGCAAGGCTGTGCTGCCACCCAGATCTTCTTCAAGTGCATGCGGNA 240
QY 327 TGGTGGCCAGGTGGACCACTGATCGAGCGCAGCAGACGCGGAGCAGTGGTGAGCGG 386

241 TGGGTCGGCAGGTCGATACGTAATTTGAACGAGCAGCAGCAGCGGATTCGCA 300
387 ACCATGGAGTCGGCAATTAATGCTCAAGGACAGAGTCAGAAAGGTGTTGGGCTGGAG 446
301 ACCATGGAGTCCGTATCAACGTATCGAGGACAGGTAGAGAAAGTGTGGGCTGGAG 360
447 GAATATGTTGCCAGTGTGGATGGAGTCTCTGGGCAACCAACGAAGTCAAGAGGCC 506
361 GAATATGTTGCCAGTGTGGATGGAGTCTCTGGGCAACCAACGAAGTCAAGAGGCC 420
507 TACATCGCATAGCCAGATTCGCCAAGACTTTCAGCAGGAGCGGTGGCCAAAGGAGGTG 566
421 TACATCGCGATTCGCTATTTGGCCAGAAGACTTTAGCAGGAGCACTGGCCAAAGGAGTG 480
567 TCCTCTGAGCAGGCGATCGTGTGCTGCTGCCGTCCACCGGAGGCGATCCCTCCAGCGGAG 626
481 TCACTGGAGCAAGCAATGTACTACTTTGTGCGCCGCCAGAGGAATCCGCCAGCTGAG 540
627 GTGGAGTGGTCCGGAAAGGAGACTGGTGGACCGGTCCCTGGACCCCAATGTATATC 686
541 GTGGAGTGGTTCGAAATAGGAGACTCTGTTGAGACCCCTCCCTCGATCCCAATGTGTATC 600
687 ACGCGGAGCACAGCTGTGTGTCGACAGCGCGCTTGTCTGACACGGCCAACTACAC 746
601 ACGCGGAGCACAGCTGTGTGTCGACAGCGCGCTTGTCTGACACGGCCAACTACAC 660
747 TGGTGGCCCAAGAACATCGTGGCACTGCGCGCAGCGCTCCGCTGTCTGTCTATGCTATC 806
661 TGTGTGCCCAAGAACATCGTAGCCCGTGGCGGAAGCACCTCTGCGAGCGGTCAITGTTAT 720
807 GTGAACCGTGGTGGTTCGAGTGGACAGAGTGTCTGCTGTGACGCGCAGCTGTGGCGC 866
721 GTGAACCGTGGTGGTTCGAGTGGACAGAGTGTCTGCTGTGACGCGCAGCTGTGGCGC 780
867 GGCTGGCAGAAACGAGCGAGTGTGACCAACCCGCGCGCTCTCAACGGGGCGCTTTC 926
781 GGCTGGCAGAAACGAGCGAGTGTGACCAACCCGCGCGCTCTCAACGGGGCGCTTTC 840
927 TGTGAGGGGAGAAATGTCATAGCCGACCGCTCTCTCTCTGTGTGTCTGTGTGGACGC 986
841 TGTGAGGGGAGAAATGTCATAGCCGACCGCTCTCTCTCTGTGTGTCTGTGTGGACGC 897
987 AGCTGGAGCCGTGGAGGAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 1046
898 AGCTGGAGTTCGTGGAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 957
1047 CGTCAGTGTCTGACCCAGCAGCCCGCAACGAGGAGGAGTGCACGGGCACTGACCTG 1106
958 CGCGAGTGTCTGACCCAGCAGCCCGCAATGGAGGTGAGGAGTGTGGGGTGTGTGACCTG 1017
1107 GACACCCGCAACTGTACAGTGAACCTGTGTGTGACACAGTGTGTGGCCCTGTGAGGAGTG 1166
1018 GACACCCGCAACTGTACAGTGAACCTGTGTGTGACACAGTGTGTGGCCCTGTGAGGAGTG 1077
1167 GCGCTGTATGTGGCTGTATGCGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 1226
1078 GCTCTCTACATCGGCTGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 1137
1227 ATCTCTGTTATTCGCGAAGAGGAGGCTGGACTCAGATGTGGTGTGACTCGTCCATTT 1286
1138 GGACTCATTTACTGTGCAAGAAGAGGAGGCTGGACTCCGATGTGGCGACTCGTCCATC 1197
1287 CTCACCTCAGGCTTCAGCCCGTGTAGATCAAGCCCGAGCAAGCAGACCAACCCCAATCTG 1346
1198 CTCACCTCAGGCTTCAGGCTGTGAGCATCAAGCCCGAGCAAGCAGACCAACCCCACTG 1257
1347 CTCACCTCAGGCTTCAGGCTGTGAGCATCAAGCCCGAGCAAGCAGACCAACCCCACTG 1403
1258 CTCACCTCAGGCTTCAGGCTGTGAGCATCAAGCCCGAGCAAGCAGACCAACCCCACTG 1317
1404 CGGAGGATGGCCCGAGCCCAAGTTCAGTCTCAACATGGGCACTGCTCAGGCCCTG 1463

1318 AGGAGGATGAGCAGCCAGCCCAAGTTCCAGCTCTCTAATGGTCACTGTCTAGCCCACTG 1377
1464 GGTGGGGCGGCCACACACTGACACAGCTCTCCAGCTCTGAGCCGAGGAGTTCGTC 1523
1378 GGGAGTGGCGCCCATAGTTGACCAAGTCAACCACTCTGAGGCTGAGGACTTCGTC 1437
1524 TCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTCCCGAGGCAACCAACATGACC 1583
1438 TCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTCCCGAGGCAACCAACATGACC 1497
1584 TATGGACCTTCAATCTCTCCGCGGCGCTGATGATCCCTAATACAGGATACGCTC 1643
1498 TACGGACCTTCAATCTCTCCGCGGCGCTGATGATCCCTAATACAGGAGTACGCTC 1557
1644 CTATCCCGCAGATGCCATACCCGAGGAGATCTATGAGATCTACCTTCACTGAC 1703
1558 CTATACCCCGGATGCCATCCCGAGGAAAGATCTACGAGATCTACCTTCACTGAC 1617
1704 AAGCCGAAGACCTGAGGTTGCCCTTAGTGGCTGTGAGACCCCTGTGAGTCCCATCGTT 1763
1618 AAGCCAGAGAGCTGAGGTTGCCCTTAGTGGCTGTGAGACCCCTGTGAGTCCCATCGTT 1677
1764 AGCTGTGAGACCCCTGTGCTCAACCGGCGAGTCACTGTGGCTATGAGCACTGT 1823
1678 AGCTGTGGGCGCCAGAGTCTGTCTACCCGCGCAGTCACTCTGCAATGAGCACTGT 1737
1824 GGGAGCCAGCCCTGACAGCTGGAGCTCGGCTCAAAAGAGAGTCTGCGAGGGCAGC 1883
1738 GAGAGCCAGCCCTGACAGCTGGAGTCTGGCTCAAAAGAGAGTCTGCGAGGGCAGT 1797
1884 TGGAGCAGGATGTGTGCACTTGGGCGAGGAGCGCCCTCCACCTCTACTACTGCGAG 1943
1798 TGGG---AGGATGTGTGCACTTGGTGGAGAGTCACTTCCCACTCTACTACTGCGAG 1854
1944 CTGGAGCCAGTGTCTGCTAGCTTTCACGAGAGTGTGGCGGCTTTCGCTGTGGGGA 2003
1855 CTGGAGCCGCGGCTGTCTATGCTTTCAGGAGCAGTGTGGCGGCTTTCGCTGTGAGA 1914
2004 GAGCCCTCAGCTGTGGCTGCGCCAGCGCTCAAGCTGTCTGTTTGGCCGCTGGCC 2063
1915 GAGCCCTCAGCTGTGGCTGCGCCAGCGCTCAAGCGCTCTCTGTTTGGCCGCTGGCC 1974
2064 TGCACCTTCCCTCGAGTCAACATCCCGGCTTACTGCTGTGATGACACCCAGATGCACTC 2123
1975 TGTACGTCCTTGTGATACACATCCGAGTGTACTGCTTACAGACACCCAGAGCTCTC 2034
2124 AAGAGTGTGTGAGTGTGAGAGAGTGTGGGCGAGCTGTGATCCAGAGGACCGGGTC 2183
2035 AAGAGTGTGTGAGTGTGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2094
2184 CTGCACCTCAAGGACAGTATACCAACCTGCGCTATCCATCCAGATGTGCCAGCTCC 2243
2095 CTGCACCTCAAGGACAGTATACCAACCTGCGCTATCCATCCAGAGTGTGCCAGCTCC 2154
2244 CTGTGGAAGAGTAAAGCTTCTGTGAGTACAGGAGTCCCTTTTATCATCTGGAAT 2303
2155 CTGTGGAAGAGTAAAGCTTCTGTGAGTACAGGAGTCCCTTTTATCATCTGGAAT 2214
2304 GGCACGAGCGGTACTTGTGCACTGTCACTTCACTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2363
2215 GGCACGAGCGGTACTTGTGCACTGTCACTTCACTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2274
2364 GACCTGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2423
2275 GACCTGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2334
2424 AACTTCAACATCAACAGGACACAAGGTTGTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2483
2335 AACTTCAACATCAACAGGACACAAGGTTGTGCTGAATTTGTGCTGAGTGTGAGTGTGAGTGTG 2394
2484 GGGTTCACGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2543
2395 GGGTTCACGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2454

1464 GGTGGGGGCGGCAACACATGACACAGCTCTCCACCTCTGAGGCCGAGGATTCGTC 1523
1378 GGAGTGGCGGCATACGTTGACACACAGCTACACCACTCTGAGGCTGAGGATTCGTC 1437
1524 TCCGCGCTCTCCACCCAGACTACTTCGCTCCCTGCGCGGAGGACACAGCAATGACC 1583
1438 TCCGCGCTCTCCACCCAACTACTTCGTTCCCTGCGCGGAGGACACAGCAATGAGCC 1497
1584 TATGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACAGGATATCAGCCTC 1643
1498 TACGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACAGGATATCAGCCTC 1557
1644 CTATCCCCCAGATGACATACCCGAGGAGAGATCTATGAGATCTACCTCAGCTGCAC 1703
1558 CTATACCCCGGATGACATCCCGGAGGAGAGATCTACGAGATCTACCTCAGCTGCAC 1617
1704 AAGCGGAGAGAGCTGAGTTGCCCTAGCTGGCTGTGACAGCCCTGCTGAGTCCCATCGTT 1763
1618 AAGCGGAGAGAGCTGAGTTGCCCTAGCTGGCTGTGACAGCCCTGCTGAGTCCCATCGTT 1677
1764 AGCTGTGGACCCCTGCGCTCTGCTCACCGGCGAGTCTATCTGGCTATGGACCACTGT 1823
1678 AGCTGTGGGCCCCAGAGTCTGCTCACCGGCGAGTCTATCTGGCTATGGACCACTGT 1737
1824 GGGAGCCCGGCTGACAGCTGGAGCTGGGCTCAAAAAGAGTCTGCGAGGGGAGC 1883
1738 GGAGAGCCCGGCTGACAGCTGGAGTCTGGGCTCAAAAAGAGTCTGCGAGGGGAGT 1797
1884 TGGAGAGAGAGTGTGCTGACCTGGGCGAGGAGCGGCTCCCACTCTACTACTGCGAG 1943
1798 TGGG--AGGATGTGCTGACCTTGGTGAGAGTCACTTCCCACTCTACTACTGCGAG 1854
1944 CTGAGGCGCAGTGTGCTGCTGCTTCCAGGAGAGTGGGCGGCTTTGGCCCTGGTGGGA 2003
1855 CTGAGGCGGCGGCTGCTGCTTCTTCCAGGAGAGTGGGCGGCTTTGGCCCTGGTGGGA 1914
2004 GAGCCCTCAGCGTGGCTGCGGCGCAAGGCTCAAGCTGCTTCTGTTGGCGCGGTGGCC 2063
1915 GAGCCCTCAGCGTGGCTGCGGCGCAAGGCTCAAGGCTGCTTCTGTTGGCGCGGTGGCC 1974
2064 TGCACTCCCTCGAGTCAACATCGGCTCTACTGCTGATGATGACACCCAGATGCACTC 2123
1975 TGTACGCTCCCTTGAGTACAAATCGAGTGTACTGCTTCCACGACACCCAGAGCTCTC 2034
2124 AAGGAGTGTGTGAGCTGGAAGAGAGTGGGCGGAGAGTGTATCCAGGAGCCAGCGTCTC 2183
2035 AAGGAGTGTGTGAGCTGGAAGAGAGTGTATCCAGGAGCCAGCTGATCCAGGAGCCCTCGGCTC 2094
2184 CTGCACTTCAAGGAGAGTGTACCACTGCGGCTATCCATCCAGATGTGCCAGCTCC 2243
2095 CTGCACTTCAAGGAGAGTGTACCACTGCGGCTATCCATCCAGATGTGCCAGCTCC 2154
2244 CTGTGGAAGAGTGTGAGCTGTGAGTGTGAGGAGATCCCTTTTATCACTCTGGAAT 2303
2155 CTGTGGAAGAGTGTGAGCTGTGAGTGTGAGGAGATCCCTTTTATCACTCTGGAAT 2214
2304 GGCACGAGCGGTATCTGCACTGTGAGTGTGAGTGTGAGGAGATCCCTTTTATCACTCTGGAAT 2363
2215 GGCACGAGCGGTATCTGCACTGTGAGTGTGAGTGTGAGGAGATCCCTTTTATCACTCTGGAAT 2274
2364 GACCTGGCTTCAAGCTGTGGTGTGAGTGTGAGGAGATCCCTTTTATCACTCTGGAAT 2423
2275 GACCTGGCTTCAAGCTGTGGTGTGAGTGTGAGGAGATCCCTTTTATCACTCTGGAAT 2334
2424 AACTTCAACATCAACAGAGATCAAGGATTTGTGAGTGTGAGGAGATCCCTTTTATCACTCTGGAAT 2483
2335 AACTTCAACATCAACAGAGATCAAGGATTTGTGAGTGTGAGGAGATCCCTTTTATCACTCTGGAAT 2394
2484 GGGGTCCAGCGCTGTGGTGGGCGGAGTGTGAGTGTGAGGAGATCCCTTTTATCACTCTGGAAT 2543
2395 GGGGTCCAGCGCTGTGGTGGGCGGAGTGTGAGTGTGAGGAGATCCCTTTTATCACTCTGGAAT 2454

2544 ATAAATTTTCAGCTGTGACCCACCTCTAGCGGGGTGCGGAGTCTGGGCCAG 2603
2455 ATCATCGCCAGTCTGACCCACCTCTGACCGGGGCGGAGTCTGAGGCCAG 2514
2604 AAATCTCACTGTGAGAGCCATCTCAGCTCTTTTGTCTTCAAGCCGAGCCGAGCCATG 2663
2515 AAATCTCACTGTGAGAGCCATCTCAGCTCTTTTGTCTTCAAGCCGAGCCGAGCCATG 2574
2664 ATCTCACTGTGAGAGCCGAGCTTCCCAAGCGGAACTCTCAGCAGCTGGCTGCA 2723
2575 ATCTCACTGTGAGAGCCGAGCTTCCCAAGCGGAACTCTCAGCAGCTGGCTGCA 2634
2724 GCACTGGCTGAGCTGGGCGGAGCAGCTGGCTCTTTCACAGTGTGGAGGCTGAGTGC 2783
2635 GCTGTGGCGGAGTGGGCGGAGCAGCTGGCTCTTTCACAGTGTGGAGGCTGAGTGC 2694
2784 TGA 2786
2695 TGA 2697

RESULT 8
RNU87305
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Rattus norvegicus (Norway rat)
Rattus norvegicus
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

1 (bases 1 to 2697)
Leonardo, E.D., Hincik, L., Masu, M., Keino-Masu, K., Ackerman, S.L. and
Tessier-Lavigne, M.
Vertebrate homologues of C. elegans UNC-5 are candidate netrin
receptors
Nature 386 (6627), 833-838 (1997)
97271897
9126742
2 (bases 1 to 2697)
Leonardo, E.D., Hincik, L., Masu, M., Keino-Masu, K. and
Tessier-Lavigne, M.
Direct Submission
Submitted (28-JAN-1997) Anatomy, UCSF, 513 Parnassus, San
Francisco, CA 94143-0452, USA
Location/Qualifiers
1. 2697
/organism="Rattus norvegicus"
/mol_type="mrna"
/db_xref="taxon:10116"
/tissue_type="brain and ventral spinal cord"
/dev_stage="18 day embryo and 13 day embryo"
1. 2697
/codon_start=1
/product="transmembrane receptor UNC5H1"
/protein_id="AAB57678.1"
/db_xref="GI:2055392"

CDS

/translation="MAVRPGLWPLVLLGIVLAALWLRGSGAQSATVAVPVCANPDLLP
HFLPEPVEDVIVKNPVLLVCKAVPATQIFKNGEWVRQVDHVIERTSDSSGLPTM
SLEINVRQVQEVKFLGEEYWCQCVAMSSGTTKSKAYIRIAYLRKNFEGEPLAKEV
EVGGIVLPKRPPEGIPAEVWLRNEDLVDPDNPVYITREHSLVVRQARLADTAN
YTCVAKNIVARRRSTSAIVVNGWSMTWESVCSACGRGQKRSCTNPAPLN
GGACCEQNVQKTACATLCPVDGWSWSKWSACGLDCTHWRSECSDPAPNGGEEC
RGALDTRNCTSDLCITASCPEVALYGLVAVAVCLFLLLLGLLYCRKKEGLDS
DVADSSILTSGFQVSIKPSKADNPHELLIQTDLSTTTTITGSGLSRQDGPSPFQL
SNHLLSPGLSGRTHLHSSSEADPVSLSLSTNYSPRSLPRGTNNMAYGTNPLGG
RLMPTNTGISLLIPDAIPRGYIEIYLTLPKEDVRLPLAGCQTLSPVVSQGPVG
LLTPVILAMHDCGEPSPDSRLKXKSCGSEWEDVHLGEEPSHLYYQLRAGAC
YVTFEQLGRPALVGEALSVATKRLRLLEPAVACTSLEYNIYVYCLDHTDALKEV
QLEKQLGQLIQEPRVLFHFKDSYHNLRSLSHDVPSSLWSKSLYSQIEIPVHTWGT
QQYLHCTFTLERINASTDSLACKVWVWQVGGGQSFNINFTKDTTRFAELLALESEG

Db 2035 AAGGAGTGGTGCAGCTGGAGAGCAGTAGTGGACAGCTGATCAGGAGCCTCGCGTC 2094

QY 2184 CTGCACCTCAAGACAGATTACCAACACCTGGCGCTATCCATCAACGATGTGCCAGCTCC 2243

Db 2095 CTGCACCTCAAGACAGATTACCAACACCTAGCTCTCCATCCAGCAGTGGCCAGCTCC 2154

QY 2244 CTGTGAAGAGTAAGTCTCTGTGCTAGCTACCAAGAGATCCCTTTTATCACATCTCGAAT 2303

Db 2155 CTGTGAAGAGTAAGTCTCTGTGCTAGCTACCAAGAGATCCCTTTTATCACATCTCGAAT 2214

QY 2304 GGCACGACGGGTACTTGGCACTGCACCTTCCAGCGGTGTGAGCGGCTGAGCGGCTAGT 2363

Db 2215 GGCACGACGGGTACTTGGCACTGCACCTTCCAGCGGTGTGAGCGGCTGAGCGGCTAGT 2274

QY 2364 GACCTGGCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGAGCGGCGAGCTTCAGCATC 2423

Db 2275 GACCTGGCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGAGCGGCGAGCTTCAGCATC 2334

QY 2424 AACTTCAACATCACCAGGACACAGGTTTGTGCTGAGCTGCTGAGAGTGAAGCG 2483

Db 2335 AACTTCAACATCACCAGGACACAGGTTTGTGCTGAGCTGCTGAGAGTGAAGCG 2394

QY 2484 GGGGTCCAGCCTGTGGGCGCCAGTGCCTTCAAGATCCCTTCCATTCGGCGAAG 2543

Db 2395 GGGGTCCAGCCTGTGGGCGCCAGTGCCTTCAAGATCCCTTCCATTCGGCGAAG 2454

QY 2544 ATAATTTCCAGCTGGACCCACCTGTGAGCGGGGTGCCAGTGGCGGAGCTTGGGCCAG 2603

Db 2455 ATCATCGCCAGTGTGGACCCACCTGTGAGCGGGGTGCCAGTGGCGGAGCTTGGGCCAG 2514

QY 2604 AACTTCAACATCACCAGGACACAGGTTTGTGCTGAGCTGCTGAGAGTGAAGCG 2663

Db 2515 AACTTCAACATCACCAGGACACAGGTTTGTGCTGAGCTGCTGAGAGTGAAGCG 2574

QY 2664 ATCTCAACCTGTGGGAGCGCGCACTTCCCAACGGCAACCTTCAGCCAGCTGGCTGCA 2723

Db 2575 ATCTCAACCTGTGGGAGCGCGCACTTCCCAACGGCAACCTTCAGCCAGCTGGCTGCA 2634

QY 2724 GCAGTGGCTGGAGTGGGCGAGCCAGAGCTGGCTTTCACAGTGTGGAGGCTGAGTGC 2783

Db 2635 GCTGTGGCGGAGTGGGCGAGCCAGAGTGGCTTTCACAGTGTGGAGGCTGAGTGC 2694

QY 2784 TGA 2786

Db 2695 TGA 2697

RESULT 9

BC058084

LOCUS

DEFINITION Mus musculus unc-5 homolog A (C. elegans), mRNA (cdna clone

MG:66671 IMAGE:6813463), complete cds.

ACCESSION BC058084

VERSION BC058084.1 GI:34784158

KEYWORDS MSC.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scturognathi; Muridae; Murinae; Mus. 1 (bases 1 to 3844)

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenman, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, W.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 3844)

Strausberg, R.

Direct Submission

Submitted (08-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library Preparation: M. Bento Soares, University of Iowa

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada

info@cgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Helao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>

Series: IRAC Plate: 126 Row: b Column: 11

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23346570.

FEATURES

source

1. 3844

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="MGC:66671 IMAGE:6813463"

/tissue_type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc"

/clone_lib="NIH_BMAP_FW0"

/lab_host="DH10B"

/note="Vector: pYX-ASC"

1. 3844

/gene="Unc5a"

/db_xref="LocusID:107448"

/db_xref="MGI:894682"

252. 2780

/codon_start=1

/product="Unc5a protein"

/protein_id="AAH58084.1"

/db_xref="GI:34784159"

/db_xref="LocusID:107448"

/translation="MAVRPGLMPALLGLVLTAMLRGSAQOASATVANPVGANPDLPL HFVPEPDVIVKNKPVLLVLCVQATQIFKCNQWVRQDVHVRSTDSGSLPTN EVRINVSQKQEVKFGLEEVWCOCVWSSGTTKSKAYIRIAYLRKNFEQELAKEP ITCGVILPCRPPEGIIPAEVWELNEDLVDPDLDPNVYITREHSLVVRQARLADTPN ITCVAKNIVARRSASAAVIVYDGSNPSKWSACGLDCTHWSRSCSDPAPRNGGE ECGADLTNCTSDLCHTSSGPDVALYIGLVAVACILLLILLVLIYCRKKEGL DSDVADSLTSGFPVSIKPSKADNPHTLTQDLSLTITTYQSGLCRQDGPSPKF QLSNGHILSPGSGRHTLHSSPTSEAEFDVSRSTQNYFRSLPRGTSNMAYGTFNFI GRLIMPTGISLLIPDAPRGKIYEYTLTHKPEDVRLPLAGCTQLSLHYCOLEAG GVLLTRVILLAMDHCGSPSPDSWSLRLLKQSCGSEWEDVLHGLHESPSLYCOLEAG ACVYTESQLGRFALVCGSALSVAKYRLRLLLPAPVACTSLFVNIYVYCLHDTQALKE VVQLEKQIGGQLIQEPKRVLPKDSIHNLRSLSIHDPVPSLWSKLLVSYQELPFFHWN

```
GTQVYLHCTFTLERVNSTSLACKVWVQVGDGQSFNFINITKDRFAEMLALES
EGYPAIVGSPAFKIPILIRQIIITSDPPGSGADWRTLAKLHLDLSLFFASKPS
PTMILNLWEARHPNGNLGALAAVAGLQDPDAGLFTVSEAC"
714. .926
/misc_feature
/db_xref="CDD:smart00409"
/misc_feature
84. .1130
/note="TSPI; Region: Thrombospondin type 1 repeats"
/db_xref="CDD:smart00209"
/misc_feature
1566. .1877
/note="ZU5; Region: Domain present in ZO-1 and Unc5-like
netrin receptors"
/db_xref="CDD:smart00218"
/misc_feature
2496. .2684
/note="DEATH; Region: DEATH domain, found in proteins
involved in cell death (apoptosis). Alpha-helical domain
present in a variety of proteins with apoptotic functions.
Some (but not all) of these domains form homotypic and
heterotypic dimers"
/db_xref="CDD:smart00005"

ORIGIN
Query Match 70.1%; Score 2020.6; DB 10; Length 3844;
Best Local Similarity 84.1%; Pred.No.0;
Matches 2429; Conservative 0; Mismatches 274; Indels 184; Gaps 6;

QY 1 AGCTGGGGCTCCGGGCTCAGCGCT-AAAGCCGCCCTCCGCCCGGGGCGCCGCCGCC 59
Db |||||
QY 60 GGGCGGCGCCCTGCGCGCGGGCCATGGCGTCCGGCCCGGCCCTGTGGCCAGCGCTC 119
Db |||||
QY 120 CTGGGATAGTCTCGCGCTCGCTCCGCGCTCGGGTCCGAGAGTGCACCGCTG 179
Db |||||
QY 285 CTGGGATAGTCTCACTGCTGGCTTCGGTTCGGGTCCGAGAGTGCACAGTG 344
QY 180 GCCAACCCAGTGGTGTGCAACCCGAGCTGTCTCCCACTTCTCTGTGGAGCCGCGAG 239
Db |||||
QY 345 GCCAACCCAGTGGTGTGCAACCCGAGCTGTCTCCCACTTCTCTGTGGAGCCGCGAG 404
QY 240 GATGTGATCATCTGAAGAACAGCCAGTGTCTGTGTGCAAGGCCGCTGCCCGCACG 299
Db |||||
QY 405 GAGTGTATCATTTCAAGAACAGCCGCTGCTGTGTGCAAGGCTGTGCCCGCAC 464
QY 300 CAGATCTTCTCAAGTGCAACCGGGAGTGGTTCGCGCAGTGGACCAAGTATCGAGCGC 359
Db |||||
QY 465 CAGATCTTCTCAAGTGCAACCGGGAGTGGTTCGCGCAGTTCGATCACGTCATTGAACGC 524
QY 360 AGCAGAGCGGAGCAGTGTGAGCCGACCATGGAGTCCGCATTAAATGTCTCAAGGCAG 419
Db |||||
QY 525 AGCACTGACGGCAGCAGCGGATTGCCAACATGGAGGTCCGATCAACGTATCAAGGCAG 584
QY 420 CAGGTGCAAGAGTGTTCGGGTGGAAGAACTAGTGTGCAAGTGGTGGCATGGAGCTCC 479
Db |||||
QY 585 CAGGTGCAAGAAAGTGTTCGGGTGGAAGAGTACTGGTGCAGTGTGTGGCATGGAGCTCC 644
QY 480 TCGGGCACCACCAAGAGTCAGAGGCCTACATCCGATAGCCAGATGGCGAAGAACTTC 539
Db |||||
QY 645 TCAGNACCACCAAGCAGAGGCCTACATCCGGATTGCCATTATTTGGGCAAGAACTTT 704
QY 540 GAGCAGGAGCCGCTGGCCCAAGAGGTGTCTCTGGAGCAGGCATCGTGTGCCCTGCCGT 599
Db |||||
QY 705 GAGCAGGAGCCGCTGGCCCAAGAGTGTCTCTGGAGCAAGGCATTGTGTACTTGTGCGC 764
QY 600 CCACGGAGGGCATCTCTCCAGCCGAGTGGAGTGGCTCCGGAAACGAGGACCTGTGTGAC 659
Db |||||
QY 765 CCCCAGGAAGGAATCCCCCAGCTGAGGTGGAGTGGCTCCGAAATGAGGACCTGTGTGAC 824
QY 660 CGGTCTCTGAGCCCAATGTATACATCACCGGGAGCAGCGCTGTGTGTGCGAGCGCC 719
Db |||||
QY 825 CCTCTCCCTCGACCCCAATGTATACATCACCGGGAGCAGCGCTAGTCTGTGGCGAGGCC 884
```

```
QY 720 GGCCTTGCTGACACGGCCAACTACACTGCTGTGCGCAAGAACTCGTGGCAGCTGCCGC 779
Db |||||
QY 885 GGCCTGGCCGACACTCCCAACTACACTGCTGTGCGCAAGAACTCGTGGCCGCTGCCGGA 944
QY 780 AGCGCTCCGCTGCTGTCATCTACTGTGAACGGTGGTGGTGTGACGTGGACCGAGTGG 839
Db |||||
QY 945 AGCGCTCTGGCGCGCTCA----- 963
QY 840 TCCGTCTGACAGCGCCAGCTGTGGCGCGGCTGGCAGAAACGGAGCGGAGCTGCACCAAC 899
Db ----- 963
QY 900 CCGCGCGCTCTCAACGGGGGGGCTTTCTGTGAGGGGCGAATGTCATGACCGCACCGTC 959
Db ----- 963
QY 960 TCCTCTCTGCTGTGCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCCGCCCTGT 1019
Db |||||
QY 964 -----TTGTTTATGAGATGGAGCTGGAGCCCATGGAGTAAGTGGTTCAGSCCTGC 1013
QY 1020 GGGCTGACTGCAACCCACTGCGGAGCCGTGAGTGTCTGACCCAGCACCCCGCAACGGA 1079
Db |||||
QY 1014 GGGCTTGACTGCAACCCACTGCGGAGCCGGAGTGTCTCGACCCAGCGCCCGCAACGGA 1073
QY 1080 GGGGAGAGTGCAGGGCACTGACTGGACACCCGCAACTGTACCAAGTACCTTCTGTGTA 1139
Db |||||
QY 1074 GGTGAGAGTGCAGGGGTGTGACTGGACACCCGCAACTGTACCAAGTACCTTCTGTGCTG 1133
QY 1140 CACAGTGTCTTCCGCTGAGGAGCTGGCCCTCTATGTGGGCTCATCGCGCTGTGGCGCTG 1199
Db |||||
QY 1134 CACACCTTTCGCGCCCGGAGAGCTGGTCTCTACATGGGCTCTGTGGCGCTGTGGCGCTG 1193
QY 1200 TGCTGTCTGTGCTGTGCTGTCTCATCTCTGTTTATGCGCGAAGAGGAGGGCTG 1259
Db |||||
QY 1194 TGCTCATCTGTGCTGTGCTGTGCTCTCATCTACTGCGCGAAGAGGAGGACTG 1253
QY 1260 GACTGAGTGTGGCTGACTGCTCATCTCTCAGGCTTCCAGGCTTCCAGGCTGTGAGCATCAAG 1319
Db |||||
QY 1254 GACTCAGAGCTGGCTGACTCATCTCTTACCTCAGGCTTCCAGGCTGTGAGCATCAAG 1313
QY 1320 CCAGCAAGAGCAGAACCCCATCTGCTCACCATCCAGCGCGAGACCTCAGCACCAAC-- 1376
Db |||||
QY 1314 CCAGCAAGAGCAGCAATCCCATCTGCTCACCATCCAAACCGGAGCTCAGCACCAACAG 1373
QY 1377 ACCACTTACAGGGCAGTCTCTGTCCCGCGCAGATGGGCGCCAGCCCAAGTTCAGCTC 1436
Db |||||
QY 1374 ACCACTTACAGGGCAGCTGTGTCCCGCGCAGATGGACCCAGCCCAAGTTCAGCTC 1433
QY 1437 ACCAATGGGCACTGTCTAGCCCCCTGGGTGGCGCGCCACACTGCAACCAAGCTCT 1496
Db |||||
QY 1434 TCTAATGGTCAGCTGTCTAGCCCCACTGGGCGAGTGGCGGCATACGCTGCACCAAGCTCC 1493
QY 1497 CCACCTCTGAGCGCGAGGAGTTCGCTCCCGCTCTCCAGCCAGAACTACTTCCGCTCC 1556
Db |||||
QY 1494 CCACCTCTGAGGCTGAGGACTTCGCTCCCGCTCTCCACCCAAACTACTTTCGTTCT 1553
QY 1557 CTGCCCCGAGGACCAAGCAATGACCTATGGGACCTTCAACTTCTCTCGGGCGCGGCTG 1616
Db |||||
QY 1554 CTGCCCCGCGGTACCAAGCAATGACCTATGGGACCTTCAACTTCTCTCGGGCGCGGCTG 1613
QY 1617 ATGATCCCTATACAGGTATCAGCTCTCTATCCCCCAGATGCCATACCCCGAGGGAAG 1676
Db |||||
QY 1614 ATGATCCCTAACACAGGAATCAGCTCTCTATACCCCGGAGCGCCATCCCGCGAGGAAG 1673
QY 1677 ATCTATGAGTCTACTCAGCTGCACAGCGGAGAGCTGAGGTGCCCCCTAGCTGCG 1736
Db |||||
QY 1674 ATCTAGAGATCTACTCTCACTGTGCAAGCCAGAGAGCTGAGGTGCCCCCTAGCTGCG 1733
QY 1737 TGTACAGCCCTGCTGAGTCCCATCGTTAGCTGTGACCCCTTGGCGTCTCTGTCTCACCCGG 1796
Db |||||
QY 1734 TGTACAGCCCTGCTGAGTCTCTATCGTTAGCTGTGGGCCCCCGAGGAGTCTCTGTCTACCCGG 1793
```

QY	1797	CCAGTCATCTGGCTATGACCACTGTGGGAGCCAGCCCTGCACAGCTGGAGCTGGCG	1856
Db	1794	CCAGTCATCTTGGCTATGACCACTGTGGGAGCCAGCCAGTCCGACAGCTGGAGCTGGCG	1853
QY	1857	CTCAAAAGCAGTCGTGGCAGGCGACCTGGAGCAGATGTGCTGCACCTGGCGGAGGAG	1916
Db	1854	CTCAAAAGCAGTCGTGGGAGCAGCTGGG---AGGACGTGCTGCACCTGTGTGAGGAG	1910
QY	1917	GGCCCTCCCACTCTACTACTCCAGCTGGAGGCCAGTGCCTGTACCTTTTACCAGAG	1976
Db	1911	TGCGCCCTCTCATCTCTACTACTCCAGCTGGAGGCCAGTGTATGTCTTTTACCAGAG	1970
QY	1977	CAGCTGGCGGCTTTGGCCCTGTGGAGAGGCCCTCAGCGTGCCTGCCGACAGCGCTC	2036
Db	1971	CAGCTAGGCGGCTTTGGCCCTGTGGAGAGGCCCTCAGCGTGCCTGCCAAGCGGCTC	2030
QY	2037	AAGCTGCTCTGTGTGGCGCGTGGCTGCACCTCCCTCAGGTACACATCCGGGTCTAC	2096
Db	2031	AGGCTGCTCTGTGTGGCCCTGTGGCTGTACGTCCCTCAGGTACACATCCGAGTGAC	2090
QY	2097	TGCTGCATGACACCCACGATGACCTCAAGAGGTTGGTGCAGCTGGAGAGCAGTGGGG	2156
Db	2091	TGCTGCACGACACCCACGATGCTCTCAAGAGTGGTGCAGCTGGAGAGCAGTGGGT	2150
QY	2157	GGACAGCTGATCCAGGAGCCACGGTCTCTGCACCTTCAAGACAGTTTACCAACCTCGC	2216
Db	2151	GGACAGCTGATCCAGGAGCCCGTGTCTGCACCTTCAAGACAGTTTACCAACCTACGT	2210
QY	2217	CTATCCATCCAGATGTGCCAGCTCCCTGTGAAGAGTAAAGCTCTTGTTCAGTACCAG	2276
Db	2211	CTGTCCATCCAGAGCTGCCAGCTCCCTGTGAAGAGCAGCTCTTGTTCAGTACCAG	2270
QY	2277	GAGATCCCTTTTATCATCTTGAATGGCAGCGGTTACTTGCATCGACTGCATTCACC	2336
Db	2271	GAGATCCCTTTTACCACATCTGGAATGGCACTCAGCAGTATCTGCATCGACTGCACC	2330
QY	2337	CTGGAGGCTGTGACCCACGACTAGTGTGCTGCGCTGCAGCTGTGGGTGTCAGGCTG	2396
Db	2331	CTGGAGGCGCTCAATGCACACGACGCTGCGCTGTGCAAGGTGTGGGTGTCAGGCTG	2390
QY	2397	GAGGCGACGGCAGAGCTTTCAGCATCAACTTCAACATCAACAGCAGCAGCAAGTTTGT	2456
Db	2391	GAGGAGATGACAGAGCTTCAACATCAACTTTCATCACTAAGACAGCAGGTTTGT	2450
QY	2457	GAGCTGTGCTCTGGAGATGAAGCGGGGTCCAGCCTGTGTGGGCCCCAGTGCCTTC	2516
Db	2451	GAAATGTGCTCTGGAGATGAAGGGGGGTCCAGCCTGTGTGGGCCCCAGTGCCTTC	2510
QY	2517	AAGATCCCTTCTCATTTCCGCAAGATATTTCCAGCCTGGACCCACCTGTAGCGG	2576
Db	2511	AAGATCCCTTCTCATTTCCGCAAGATATTTCCAGCCTGGACCCACCTGTAGCGG	2570
QY	2577	GGTCCGACTGGCGGACTTGGCCCCAGAACTCCACCTGGACAGCCATCTCAGCTTCTTT	2636
Db	2571	GGCGCCGACTGGCGAACTTAGCCAGAACTTCCACCTGGACAGCCATCTTAGCTTCTT	2630
QY	2637	GCCTCCAAAGCCAGCCCCACAGCATGATCTCAACTGTGGAGGCGCGCACTTCCCC	2696
Db	2631	GCCTCCAAAGCCAGCCCTACAGCATGATCTCAACTGTGGAGGCGCGCACTTCCCC	2690
QY	2697	AACGGCACTTCAGCAGCTGGCTGCAGCTGGCTGAGCTGGGCGACCCAGAGAGCTGGC	2756
Db	2691	AACGGCACTTCGCGCAGCTGGCGCAGCTGTGGCGGAGCTGTGGCCAGCAGAGTGTGGC	2750
QY	2757	CTCTTTCAGAGTGTGGAGGCTGAGTGTGTGAGGCGCGCCAGCCCGGACA--CCTACACTCT	2814
Db	2751	CTCTTTCAGGCTGAGAGGCGGAGTGTGTGAGCAGCCAGCGCGTCAAGCTCTACACTCT	2810
QY	2815	CACAGCTTTGGCACCACACAGGACAGCAGCAGCAGCAGCAGGCGGCTTCCCCACACC	2874
Db	2811	CACAGCTTTGGCACCCTCCAGGAGCAGCAGCAGCAGCAGGCGGCTTACCCCCACAC	2870
QY	2875	GGGGGAGA 2881	

Db

2871

CCGGGGA 2877

RESULT 10

BC009333

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BC009333

Homo sapiens cDNA clone IMAGE:4126760, partial cds.

2688 bp mRNA linear PRI 19-DEC-2003

BC009333

BC009333.2

GI:40226527

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2688)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Scapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalobio,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalius,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 2688)

Strausberg,R.

Direct Submission

Submitted (12-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Dec 19, 2003 this sequence version replaced gi:14424611.

Contact: MGC help desk

Email: cgaps-x@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;

Contact: nisc.mgc@nhgri.nih.gov

Web site: <http://www.nisc.nih.gov/>

Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripoop,S., Thomas,P.J., Touchman,J.W., Tsurgouon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,I., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 26 Row: g Column: 22.

Location/Qualifiers

1. 2688

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

FEATURES

source

/clone="IMAGE:4126760"		1757		QY	CATGTTAGCTGTGGACCCCTGGCGTCTCTCTCAACCGGCCAGTCATCTCTGGCTATGGA	1816
/tissue type="Brain, neuroblastoma"		601		Db	CATGTTAGCTGTGGACCCCTGGCGTCTCTCTCAACCGGCCAGTCATCTCTGGCTATGGA	660
/clone_lib="NIH MGC 19"		1817		QY	CCACTGTGGGAGCCCGCTGACGCTGAGAGCTGGCGCTCAAAAGAGAGTGTGTGGA	1876
/lab_host="DH10B-R_19"		661		Db	CCACTGTGGGAGCCCGCTGACGCTGAGAGCTGGCGCTCAAAAGAGAGTGTGTGGA	720
/note="vector: POTB7"		1877		QY	GGGAGAGCTGGAGCAGGATGTCTGACCTGGGGGAGAGGCGCTCCACCTCTACTA	1936
<1..1627		721		Db	GGGAGAGCTGGAGCAGGATGTCTGACCTGGGGGAGAGGCGCTCCACCTCTACTA	777
/codon_start=2		1937		QY	CTGCCAGCTGGAGCAGTGTCTGCTCTTCAACGAGAGCTGGGCGCTTTTGCCCT	1996
/product="Unknown (protein for IMAGE:4126760)"		778		Db	CTGCCAGCTGGAGCAGTGTCTGCTCTTCAACGAGAGCTGGGCGCTTTTGCCCT	837
/db_xref="GI:40226528"		1997		QY	GGTGGAGAGCCCTCAGCGTGGCTGGCGGAGAGCGCTCAAGTGTCTTCTGTTGGCC	2056
/translation="DVALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVDSSILT		838		Db	GGTGGAGAGCCCTCAGCGTGGCTGGCGGAGAGCGCTCAAGTGTCTTCTGTTGGCC	897
SGFQVSKADNPHLITQPLSTTTTQGSCLPQDQSPKFTQLTNGHLLSL		2057		QY	GGTGGAGAGCCCTCAGCGTGGCTGGCGGAGAGCGCTCAAGTGTCTTCTGTTGGCC	2116
GGGHTLHSPSEAEFVRSLSQNTFRSLPRGTSMNTYGTFLGRLMIPTNGI		898		Db	GGTGGAGAGCCCTCAGCGTGGCTGGCGGAGAGCGCTCAAGTGTCTTCTGTTGGCC	2116
MLHPDAIPRKIYELTLHKPEDVRLPLAGCOTLLSPVSGPGVLLTRPVILA		2117		QY	TGCACTCAAGAGAGTGTGTGAGTGGAGAGAGCTGGGGGAGAGCTGATCCAGGAGCC	2176
MDHCEPSPSWRLKQSCGSEWDLHLGERAPSHLYVCOLEASACVFFBQLGR		958		Db	TGCACTCAAGAGAGTGTGTGAGTGGAGAGAGCTGGGGGAGAGCTGATCCAGGAGCC	2176
FALVGEALSAARLKLLEFAPVACTSLEYNRYVCLHDTDLKKEVLEKLGQ		2177		QY	ACGGTCTCTGCACTTCAAGAGAGTGTGTGAGTGGAGAGAGCTGGGGGAGAGCTGATCCAGGAGCC	2236
LIOBPVHLFKDSVHNLRLSIHDVSPSSKLLVSYOEIPFFYHWNQRYLHCTFT		1018		Db	ACGGTCTCTGCACTTCAAGAGAGTGTGTGAGTGGAGAGAGCTGGGGGAGAGCTGATCCAGGAGCC	2236
LERVPSITLACKLQWQVEGQSFISFNITKOTRFAELLALAESEAGVPALVGS		2237		QY	CAGCTCCCTGTGGAAGAGTAAAGCTCCTTGTGCTACCAAGAGAGTCCCTTTTATACAT	2296
AFKIPFLIRQIISLPPRCRGADMRLEAKLHLSLSPFFASKPSPTAMILNWEA		1078		Db	CAGCTCCCTGTGGAAGAGTAAAGCTCCTTGTGCTACCAAGAGAGTCCCTTTTATACAT	2296
RHPFNGLSQLAAAVAGLQPDAGLFTVSEAC"		2297		QY	CTGGAATGGACCCAGCGGTACTTGTGCTGCACTTCAACCTGGAGCGTGTGAGCGCCAG	2356
/note="ZU5; Region: Domain present in ZO-1 and Unc5-like		1138		Db	CTGGAATGGACCCAGCGGTACTTGTGCTGCACTTCAACCTGGAGCGTGTGAGCGCCAG	2356
netrin receptors"		2357		QY	CAGTGTGACCTGGCGTCAAGCTGTGGTGTGGAGTGGAGGCGAGCGGAGAGCTT	2416
/db_xref="CDD:smart00218"		1198		Db	CAGTGTGACCTGGCGTCAAGCTGTGGTGTGGAGTGGAGGCGAGCGGAGAGCTT	2416
1343..1594		2417		QY	CAGCATCAACTTCAACATCAACAGGACACAAAGTGTGCTGAGCTGTGCTGTGAGAG	2476
/note="DEATH; Region: DEATH domain, found in proteins		1258		Db	CAGCATCAACTTCAACATCAACAGGACACAAAGTGTGCTGAGCTGTGCTGTGAGAG	2476
involved in cell death (apoptosis)"		2477		QY	TGAAGCGGGGTCCAGCCCTGGTGGGCGCCAGTGTCTTCAAGATCCCTTCTCATTCG	2536
/db_xref="CDD:smart00005"		1318		Db	TGAAGCGGGGTCCAGCCCTGGTGGGCGCCAGTGTCTTCAAGATCCCTTCTCATTCG	2536
ORIGIN		2537		QY	GCAGAGATAATTTCCAGCTGGAGCCACCTGTAGGCGGGGTGCCAGTGTGCGGAGCTCT	2596
Query Match		1378		Db	GCAGAGATAATTTCCAGCTGGAGCCACCTGTAGGCGGGGTGCCAGTGTGCGGAGCTCT	2596
Best Local Similarity 58.7%; Score 1691.4; DB 9; Length 2688;		2597		QY	GGCCAGAAACTCCACCTGGAGCGCGGCTCTTGTGCTTCCAAAGCGGAGCGCCAC	2656
Matches 1718; Conservative 0; Mismatches 1; Indels 6; Gaps 2;		1438		Db	GGCCAGAAACTCCACCTGGAGCGCGGCTCTTGTGCTTCCAAAGCGGAGCGCCAC	2656
QY 1160 GGAGCTGGCCCTCTATGTGGGCTCATCGCGGTGGCGCTGTGCTGTGCTGTGCT		2657		QY	AGCCATGATCTCAACCTGTGGAGCGCGGCTCTTGTGCTTCCAAAGCGGAGCGCCAC	2716
Db 1 GGAGCTGGCCCTCTATGTGGGCTCATCGCGGTGGCGCTGTGCTGTGCTGTGCT		1498		Db	AGCCATGATCTCAACCTGTGGAGCGCGGCTCTTGTGCTTCCAAAGCGGAGCGCCAC	2716
QY 1220 TGTCTCTCATCTCGTTTATTCGGAAGAGAGGGGCTGGATCGATGTGCTGACTC		2717		QY	GGCTCAGAGTGTGCTGGAGCGCGGCTCTTGTGCTTCCAAAGCGGAGCGCCAC	2776
Db 1 TGTCTCTCATCTCGTTTATTCGGAAGAGAGGGGCTGGATCGATGTGCTGACTC		1558		Db	GGCTCAGAGTGTGCTGGAGCGCGGCTCTTGTGCTTCCAAAGCGGAGCGCCAC	2776
QY 1280 GTCCATCTTCACTCAGGCTTCCAGCCGCTGAGCATCAAGCCAGAAAGCAGCAACCC		2777		QY	TGAGTGTCTGAGCGCGGCGGCTCTTGTGCTTCCAAAGCGGAGCGCCAC	2836
Db 121 GTCCATCTTCACTCAGGCTTCCAGCCGCTGAGCATCAAGCCAGAAAGCAGCAACCC		1618		Db	TGAGTGTCTGAGCGCGGCGGCTCTTGTGCTTCCAAAGCGGAGCGCCAC	2836
QY 1340 CCATCTGCTCACCATCCAGCCGAGCTCAG--CACCACACACCTACAGGGGAGTCT						
Db 181 CCATCTGCTCACCATCCAGCCGAGCTCAG--CACCACACACCTACAGGGGAGTCT						
QY 1397 CTGTCCCGGAGAGTGGCGCCAGCCCAAGTTCCAGCTCACCAGTGGGAGCTGCTCAG						
Db 241 CTGTCCCGGAGAGTGGCGCCAGCCCAAGTTCCAGCTCACCAGTGGGAGCTGCTCAG						
QY 1457 CCCCCTGGTGGCGCGCCACACTGACCAAGCTCTCCACCTCTGAGCGCCAGGA						
Db 301 CCCCCTGGTGGCGCGCCACACTGACCAAGCTCTCCACCTCTGAGCGCCAGGA						
QY 1517 GTTCGCTCCGCTCTCCACCGAGAACTCTTCCGCTCCCTGCGCGGAGCAGCAAA						
Db 361 GTTCGCTCCGCTCTCCACCGAGAACTCTTCCGCTCCCTGCGCGGAGCAGCAAA						
QY 1577 CATGACTATGGAGCTTCAACTCTTCCGGGCGGCTGATGATCCCTAATACAGGAT						
Db 421 CATGACTATGGAGCTTCAACTCTTCCGGGCGGCTGATGATCCCTAATACAGGAT						
QY 1637 CAGCTCTCTATCCCCCAGATGCCATACCCGAGGAGAGATCTATGATCTTACCTCAC						
Db 481 CAGCTCTCTATCCCCCAGATGCCATACCCGAGGAGAGATCTATGATCTTACCTCAC						
QY 1697 GCTGCAAGCGGAGAGAGTGGGTTGCCCTGTAGTGTGCTGTCAGACCTGTGAGTCC						
Db 541 GCTGCAAGCGGAGAGAGTGGGTTGCCCTGTAGTGTGCTGTCAGACCTGTGAGTCC						

QY 2670 AACCTGGGAGGCGCGCACTTCCCAACGGCAACCTCAGCAGCTGGCTGCAGCAGT 2729
Db 1552 AACCTGGGAGGCGCGCACTTCCCAACGGCAACCTCAGCAGCTGGCTGCAGCAGT 1611
QY 2730 GCTGAGCTGGGCGAGCAGAGCTGGGCTC-TTCACAGT-TCGGAGGCTGAGTGTGAG 2787
Db 1612 GCTGGGAGTGGCGCAGAGGAGGCTGGCTTCTTTCACAGTTCGGAGGCTGAGTGTGAG 1671
QY 2788 GCCGGCAGGCGCCACACCTACACTCTCACCAGCTTTGGCA--CCCACCAAGGACAGCA 2845
Db 1672 GCCGGCAGG--CGAACACTACAAATTTACAGTTTTGGGAACCCACCAAGGACAGCA 1729
QY 2846 GAAGCCGAGAGGGGCTTCCCAACCGGGGAGA 2881
Db 1730 GAAGCCGAGAGGGGCTTCCCAACCGGGGAGA 1765

RESULT 12
AX054976
LOCUS
DEFINITION
Sequence 91 from Patent WO0073328.
ACCESSION
AX054976
VERSION
AX054976.1 GI:12228344
KEYWORDS
synthetic construct
SOURCE
synthetic construct
ORGANISM
artificial sequences.
REFERENCE
1
AUTHORS
van Crieke, W., Roelens, I., Bogaert, T. and Verwaerde, P.
TITLE
Unc-5 constructs and screening methods
JOURNAL
Patent: WO 0073328-A 91 07-DEC-2000;
Devgen NV (BE)
FEATURES
Location/Qualifiers
source
1..9700
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="plasmid pGC1037"

ORIGIN
Query Match 44.8%; Score 1291.2; DB 6; Length 9700;
Best Local Similarity 99.5%; Pred. No. 9e-204;
Matches 1306; Conservative 0; Mismatches 3; Indels 3; Gaps 1;
QY 1478 CACACTGCACCAAGCTCTCCACCTCTGAGCGGAGGAGTTCGTCTCCGCTCTCCAC 1537
Db 7 CACACTGCACCAAGCTCTCCACCTCTGAGCGGAGGAGTTCGTCTCCGCTCTCCAC 66
QY 1538 CCAGAACTACTTCGGTCTCCCTGCGCCGAGGACCAACATGACCTATGGGACCTTCAA 1597
Db 67 CCAGAACTACTTCGGTCTCCCTGCGCCGAGGACCAACATGACCTATGGGACCTTCAA 126
QY 1598 CTTCTCGGGGGCGGCTGATGATCCCTAATACAGTATCAGCTCTATCCCCCAGAG 1657
Db 127 CTTCTCGGGGGCGGCTGATGATCCCTAATACAGTATCAGCTCTATCCCCCAGAG 186
QY 1658 TGCCATACCCCGGAGGAAGATCTATGAGATCTACCTCAGCTGCACAAACCGGAGACGT 1717
Db 187 TGCCATACCCCGGAGGAAGATCTATGAGATCTACCTCAGCTGCACAAACCGGAGACGT 246
QY 1718 GAGTTGCCCCCTAGTGGTGTGACACCTGCTGAGTCCCATCTGTTAGTGTGGACCCCC 1777
Db 247 GAGTTGCCCCCTAGTGGTGTGACACCTGCTGAGTCCCATCTGTTAGTGTGGACCCCC 306
QY 1778 TGGCGTCTGCTCACCAGGCGGAGTCTCTGCTATGAGACCACTGTGGGAGCCAGCCCC 1837
Db 307 TGGCGTCTGCTCACCAGGCGGAGTCTCTGCTATGAGACCACTGTGGGAGCCAGCCCC 366
QY 1838 TGACAGCTGAGCGCTGCGCTCAAAAAGCAGTCTGTCGAGGCGAGCTGGGACAGATGT 1897
Db 367 TGACAGCTGAGCGCTGCGCTCAAAAAGCAGTCTGTCGAGGCGAGCTGGG---AGGATGT 423
QY 1898 GCTGACCTGGGCGAGGAGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGC 1957

Db 424 GCTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACTACTGCAGCTGGAGGCGCAGTGC 483
QY 1958 CTGCTAGCTCTTCCACCAGCAGCTGGGCGGCTTTGGCCCTGTGGAGAGGCGCTCAGCGT 2017
Db 484 CTGCTAGCTCTTCCACCAGCAGCTGGGCGGCTTTGGCCCTGTGGAGAGGCGCTCAGCGT 543
QY 2018 GGCTGCGCCGCAAGCGCTCAAGCTGTCTTGTGTTGCGCGGCTGCTGCACTCCCTCGA 2077
Db 544 GGCTGCGCCGCAAGCGCTCAAGCTGTCTTGTGTTGCGCGGCTGCTGCACTCCCTCGA 603
QY 2078 GTACCAACATCGGCTCTACTGCTGCTGATGACACCCAGATGCACTCAAGAGGTTGTGCA 2137
Db 504 GTACCAACATCGGCTCTACTGCTGCTGATGACACCCAGATGCACTCAAGAGGTTGTGCA 663
QY 2138 GCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGCTCTGCACTTCAAGGA 2197
Db 664 GCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGCTCTGCACTTCAAGGA 723
QY 2198 CAGTTACCAACCTGGGCTCTATCCATCCAGATGTCAGCTCCCTGTGGAGAGTAA 2257
Db 724 CAGTTACCAACCTGGGCTCTATCCATCCAGATGTCAGCTCCCTGTGGAGAGTAA 783
QY 2258 GCTCTCTGTGCTAGCTACCGAGGATCCCTTTTATCAGATCTGGAATGGCAGCGGTA 2317
Db 784 GCTCTCTGTGCTAGCTACCGAGGATCCCTTTTATCAGATCTGGAATGGCAGCGGTA 843
QY 2318 CTTGCACTGCACCTTCAACCTGGAGGCTGTCAGCCCGCAGCACTAGTACCTGCCCTGCAA 2377
Db 844 CTTGCACTGCACCTTCAACCTGGAGGCTGTCAGCCCGCAGCACTAGTACCTGCCCTGCAA 903
QY 2378 GCTGTGGTGTGCGAGTGGAGGCGGCGAGGCTTTCAGCATCAACTTCAACATCAC 2437
Db 904 GCTGTGGTGTGCGAGTGGAGGCGGCGAGGCTTTCAGCATCAACTTCAACATCAC 963
QY 2438 CAAGGACACAAGGTTGCTGAGCTGCTGCTGAGAGTGAAGCGGGGCTCCAGCGCT 2497
Db 964 CAAGGACACAAGGTTGCTGAGCTGCTGCTGAGAGTGAAGCGGGGCTCCAGCGCT 1023
QY 2498 GGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCTCATTCGGCAGAAATATTTCCAGCCT 2557
Db 1024 GGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCTCATTCGGCAGAAATATTTCCAGCCT 1083
QY 2558 GACCCACCTGTAGCGGGGTGCGACTGGCGGACTCTGGCCCGAGAACTCCACCTGGA 2617
Db 1084 GACCCACCTGTAGCGGGGTGCGACTGGCGGACTCTGGCCCGAGAACTCCACCTGGA 1143
QY 2618 CAGCCATCTCAGCTTCTTTCCTCCAAAGCCAGCCCGCCAGCATGATCTCAACTGTG 2677
Db 1144 CAGCCATCTCAGCTTCTTTCCTCCAAAGCCCGCCAGCATGATCTCAACTGTG 1203
QY 2678 GGAGGCGCGGCACTTCCCAACGGCAACCTCAGCAGCTGGCTGAGCAGTGGCTGACT 2737
Db 1204 GGAGGCGCGGCACTTCCCAACGGCAACCTCAGCAGCTGGCTGAGCAGTGGCTGACT 1263
QY 2738 GGGCGAGCCAGACGCTGGCTCTTTCAGTGTGCGAGGCTGAGTGTGAGGC 2789
Db 1264 GGGCGAGCCAGACGCTGGCTCTTTCAGTGTGCGAGGCTGAGTGTGAGTC 1315

RESULT 13
MMU72634
LOCUS
DEFINITION
Mus musculus rostral cerebellar malformation protein (rcm) mRNA,
complete cds.
ACCESSION
U72634
VERSION
U72634.1 GI:2088526
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 9299)

Db 1679 TAGAATAGAGCCCTTAACCTGAAGAACAGAGCTCGAAGACAGACTGACCATCT 1738
QY 1576 ACATGACCTATGGGACCTTCAACTTCTCGGGGCGCGCTGATGATCCCTAATACAGTA 1635
Db 1739 GCACAGANTTTGGTACCTTCAACTTCTTGGGGGTACCTCATCATTTCTAATTCAGGAG 1798
QY 1636 TCAGCCTCTCATCCCCCAGATGCCATACCCAGGGAAGATCTATGAGATCTACCTCA 1695
Db 1799 TAAGCTGCTGATTCGCGCTGGGCCATCTCTCAGGGGAGAGTCTATGAATGATGTGA 1858
QY 1696 CGCTGCACAAGCCGAGAGAGTGAAGTTGCCCTAGCTGTGCTGTGACACCTGCTGATC 1755
Db 1859 CTGTACACAGGAAGAATAATATGAGGCCGCCATGGAAGACTCTCAGACCTCTTACCC 1918
QY 1756 CCATCGTTAGCTGTGGACCCCTCGCGTCCTGCTCACCGGCCAGTCTATCTGCTATGG 1815
Db 1919 CTGTGCTGAGCTGTGGGCTCTCTGAGCTGTGCTGACCGGCCCTGTCATCTCATCTGC 1978
QY 1816 ACCACTGTGGGAGCCAGCCCTGACAGCTGGAGCCTCGCCTCAAAAAGCAGTGTGCG 1875
Db 1979 ATCACTGTGAGACCCAGCAGCAGGACTGGAAGATCCAGCTCAAAAACAGGACGTGC 2038
QY 1876 AGGCAGCTGGGACAGAGTGTGCTGACCTGGCGAGAGGCGCCCTCCACCTCTACT 1935
Db 2039 AGGACAAATGGGAGGATGTTGTGGTG---GTTGGGAGAGAACTTCACAAACCCCTGTT 2095
QY 1936 ACTGCCAGCTGGAGGCCAGTGCCTGTCTACCTTTCACCGAGCAGCTGGGCGCTTTGCC 1995
Db 2096 ACATTAGCTGGATGAGAGCTTGCAATATCTTCACAGAAACCTCAGTACCTATGCC 2155
QY 1996 TGTGGGAGAGGCCCTCAGCTGGCTGCCGCCAGCCCTCAAGCTCTTCTGTTTGGCG 2055
Db 2156 TGGTTGGGAGTCCACCAACAAAGCAGTGCAGAGCTTTAACTGGCCATCTTTGGGC 2215
QY 2056 CGGTGCTGACCTGCTCAGTACAAATCCGGGTCTACTCGCTGATGACACCCACG 2115
Db 2216 CCTCTGCTGCTCTTCCCTGGAGTACAGCATTAGAGTCTACTGCTGGATGACACACG 2275
QY 2116 ATCACTCAAGAGGTGTGACCTGGAGAGCAGCTGGGGACAGCTGATCCAGAGC 2175
Db 2276 ATGCTCTGAGGAGTCTTCAAACTGGAGGCAATGGGAGACAGCTCTTAGAAGAAC 2335
QY 2176 CAGGCTCTGCACTTCAAGGACAGTTACCAACCTGGCCCTATCATCCAGATGTC 2235
Db 2336 CCAAGGCTCTTCAATTTAAAGGAGCATCCAACTGGCCCTGTATTATGACATCG 2395
QY 2236 CAGCTCCCTGTGGAAGATAGCTCTGTGCTGCTACAGGAGATCCCTTTTATCACA 2295
Db 2396 CCATTCCTCTGGAAGAGCAATTCGTGCTAAGTATCAGGAAATTCATTTTACCACA 2455
QY 2296 TCTGGAATGGCACGACGCGTACTTCACTGCACTTCCCTTACCTGGAGGCGGACGAGCT 2415
Db 2456 TCTGAGTGGCTCTCAAGAAACCTCACTGACCTTCACTCTGGAAGAGGCGACTC 2515
QY 2356 GCATAGTACCTGCGCTTCAAGCTGTGGGTGTGGAGTGGAGGCGGACGAGCT 2415
Db 2516 ACACAGTGAACCTGTTTCAAACTCTGTGTGGGAGCTTGAAGAGAGGCGAGATCT 2575
QY 2416 TCAGCATCACTTCAACATCAACAGGACAAGGTTTCTGAGCTGTGGCTCTGGAGA 2475
Db 2576 TCCAGCTCAACTGTACTGTGTGAGAGAACTTCTGGCATGCACTTACCTCTCTGGACC 2635
QY 2476 GTGAGCGGGGTCCAGCGCTGTGGGCCCGCAGTCCCTTCAAGATCCCTTCTCATTC 2535
Db 2636 CTGTAGTACCATCACTGATCAACGAGCAAGTGTCTTTCAGCATTCCTCTCCCTATCC 2695
QY 2536 GCGAGAAGATAATTTCCAGCTGGAGCCACCTGTAGCGGGGTGCGGAGTGGCGGACTC 2595
Db 2696 GCGAGAAGTATGACAGCCTGTGATGCCCTCAACAGAGGCGCATGATCGAGAGTGC 2755
QY 2596 TGGCCCAAGAACTCCACCTGGACAGCATCTCAGCTTCTTTCCTTCCAGCCAGCCCA 2655

Db 2756 TGGCCCATAACTCAACTGGACAGGTACTTGAATTACTTTGGCCACCAATTCGAGCCCAA 2815
QY 2656 CAGCCATGATCTCAACTCTGTGGGAGCGCGGCACTTCCCAAGCGCAACCTTCAGCCAGC 2715
Db 2816 CTGCGCTAATCTCGATCTTTGGGAAGCACAGAACTTCCAGATGGAACCTTCAGCATGC 2875
QY 2716 TGSCTGACAGAGTGGCTGGACTGGCCAGCAGCAGCTGGCTCTTCACTGTGGAGG 2775
Db 2876 TGGCAGCGCTCTGGAAGAAATGGGAAGACATGACAGTGGTGTCTCTTGGCAGCAGAAG 2935
QY 2776 CTGAGTGTCTGA 2786
Db 2936 GACAGTATTGA 2946
RESULT 14
AY187310
LOCUS AV187310 2962 bp mRNA linear VRT 06-JUN-2003
DEFINITION Gallus gallus UNC5-like protein 3 mRNA, complete cds.
ACCESSION AY187310
VERSION AV187310.1 GI:31442350
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 2962)
AUTHORS Guan,W. and Condic,M.L.
TITLE Characterization of Netrin-1, Neogenin and cUNC-5H3 expression during chick dorsal root ganglia development
JOURNAL Gene Expr. Patterns 3, 369-373 (2003)
REFERENCE 2 (bases 1 to 2962)
AUTHORS Guan,W. and Condic,M.L.
TITLE Direct Submission
JOURNAL Submitted (26-NOV-2002) Neurobiology & Anatomy, University of Utah, 20 North, 1900 East, Salt Lake City, UT 84132-3401, USA
FEATURES
source
1..2962
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
CDS
1..2796
/codon_start=1
/product="UNC5-like protein 3"
/protein_id="AA067275.1"
/db_xref="GI:31442351"
/translation="MGKLGTEAARCGLGMVLLHSHVLPALAVLGASRPGSAQQDD
FPEHLEPTFPSPPEPLPHFLIEPREAYIVKNKPNLYCKASPATQIYFKCNSEWHQ
KDHVDERVDETSGLIVCEVSIIEISRQVVEELFGPEYWCQCVAMSSAGTTSRKAYV
TIADLRKTFPEPLGKESLQEVLLQCRPEGIPIVAEVELKNEVIDPVEDRNYI
TIDHNLIIKQARLSDTANTYCVAKNIVAKRKSTATTIVVNGWSTWEMWACNSRC
GRGFORRTCTNPAPLNGAFCEGONQKIACTTLCVDGKWTSMKSTCGTCTH
WRRECTAPKNGKDCGLVLOSNCCTDGLCQAPSDVDVALIVGIVIAIVCLA
LSVVALFVYKRNHRDFRSDIIDSSALNGFPQVNIKAARDLLAVPDLTSAAMYR
GPVALHDVSKIPMNSPILDLPLNLKIVNTSGAVTPQDELDFDFSKLSPQITQS
LLENELTNVQCSLARQDTPSCFTAFNLSGHLVINSVLSLIPAGVPGQRYE
MYTVHRKEGMRPPEVDSQTLITPVSGPPGALLTRPWLTHHCAENMDMDQIOL
KHQAGQWEDVVVVEENFTTPCYIOLDPEACHILLTETLSTVALVGSITKAAAKRL
KLHGLPSCSLEYSIRVCLDQDALKVELQLEROMGQLLEPEKLPFGSTHN
LRSLHDLAHLWKSLEPAKQEIFPIHWSGQQRNLHCTFLERFSLTLELVCKLC
VROVEGEOIFOLNCVSEBPTDIPYIMDSAGSITTVGNFAFIPPIROKLGSL
DAPQTRGHRMLAHLKLDRLYLNFATKSPSTGVILDWEAQNPEDNLSMLAAVLE
EMGRHETVTVSAAAGNY"


```
Db 2363 TCAATACCTCGAGCTCGTCTGCAAACTCTGTGTGGCAAGTCGAAGGAGGCGAGA 2422
Qy 2413 GCTTCAGCATCAACTTCAACATCACCAGGACACAAGGTTTGTGAGCTGTGCTCTGG 2472
Db 2423 TCTTCCAGCTGAATCTGCTAGTATCAGAGGACCCACTGGCAATTGATTATCCCATCTGG 2482
Qy 2473 AGAGTGAAGCGGGGTCCAGCCTGTGTGGGCCCGAGTGCCTTCAAGATCCCTTCTCTCA 2532
Db 2483 ATTCAAGCAGGAGCATCACAGCATAGTTGGGCCCAAGCTTTTCAAGATCCCTTCCAA 2542
Qy 2533 TTCCGCGAAGATAATTTCCAGCTGTGACCCACCCTGTAGGCGGGTGCAGACTGGCGGA 2592
Db 2543 TAAGGCGAAGCTCTGCAAGCAGCTGTGATGACCCCAAGCCCGGGCCATGACTGGAGGA 2602
Qy 2593 CTCTGGGCCAGAACTCCACTGACAGCCATCTCAGCTTCTTTGCTTCCAGCCAGCCAGCC 2652
Db 2603 TGCTGGGCCACAGCTGAATTTGACAGGTACCTAAATTAATTTTGTACGAAGTCGAGTC 2662
Qy 2653 CCACAGCATGATCTCAACTGTGTGGAGCGCGCACTTCCCAACAGGCAACCTCAGCC 2712
Db 2663 CCACTGGGGTGTCTGATCTCTGGGAAGCCAGAAATTTCCCTGATGGCAACCTGAGCA 2722
Qy 2713 AGCTGGCTGACAGTGGCTGGACTGGCCAGCCAGCAGCTGGCTCTTCAAGTGTCTGG 2772
Db 2723 TGCTGGCAGCAGTGTCTGGAAGAAATGGGAGCATGAAGACCGTTGTTCTTTGGCAGCAG 2782
Qy 2773 AGGCTGAGTCTGAGGCGCGCCAGGCCGGA 2802
Db 2783 AAGGAATTAATGATGAGCCTTTGGCGTGA 2812
```

RESULT 15

```
AB118026
LOCUS Rattus norvegicus Unc5h3 mRNA, complete cds.
DEFINITION Rattus norvegicus Unc5h3 mRNA, complete cds.
ACCESSION AB118026
VERSION AB118026.1 GI:40217509
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1
AUTHORS Kuramoto, T., Kuwamura, M. and Serikawa, T.
TITLE Rat neurological mutations cerebellar vermis defect and hobble are caused by mutations in the netrin-1 receptor gene Unc5h3
JOURNAL Mol. Brain Res. (2003) In press
REFERENCE 2 (bases 1 to 9328)
AUTHORS Kuramoto, T. and Serikawa, T.
TITLE Direct Submission
JOURNAL Submitted (22-AUG-2003) Takashi Kuramoto, Institute of Laboratory Animals, Graduate School of Medicine, Kyoto University; Yoshidakonocho-cho, Sakyo-ku, Kyoto 606-8501, Japan (E-mail:tkuramoto@anim.med.kyoto-u.ac.jp, URL:www.anim.med.kyoto-u.ac.jp, Tel:81-75-753-4494, Fax:81-75-753-4409)
```

FEATURES

```
Location/Qualifiers
1..9328
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/chromosome="2"
gene 1..9328
/gene="Unc5h3"
CDS 151..2946
/gene="Unc5h3"
/codon_start=1
/protein_id="BAD05181.1"
/db_xref="GI:40217510"
/translation="MRKGLRATAARCGLSGLYLQMLVLPALALLSASGTSAOQDD FPELPETPPDPPEPLPHFLIEPEAYIVKKNPVNLYCKASPATQIVFKNSEWVHQ
```

```
Query Match 32.5%; Score 936.8; DB 10; Length 9328;
Best Local Similarity 61.8%; Pred. No. 3.6e-145;
Matches 1631; Conservative 0; Mismatches 937; Indels 72; Gaps 6;
Qy 204 CCGGACCTGCTCCCCACATTCCTGGTGGAGCCCGAGGATGTTACATCGTCAAGAACAAAG 263
Db 322 CCTGAGCCATTACACATTTCTTATTGAGCCCGAGGAGCTTACATTTGAGAGAACAAAG 381
Qy 264 CCAAGTGTCTTGTGTGCAAGGCCGTGCCCGCCAGCAGATCTTCTTCAAGTGCAACGGG 323
Db 382 CTTGTAACCTGTATTGTAAGCCAGCCCTGTCACCCAGATCTTCTTCAAGTGCAACAGC 441
Qy 324 GAGTGGTGGCCAGGTGGACCAAGTGTGAGCGGAGCAGACAGCGGAGCAGTGTGTAG 383
Db 442 GAATGGTTCATCAGAAGGACCAAGTAGTAGAGAGGGGTAGATGAAACCTCCGCTCTC 501
Qy 384 CCGACCATGAGGTCCGCAATTAATGTCTCAAGCAGCAGTCCGAGAGGTGTTTCGGGCTG 443
Db 502 ATTGTGAGAGAGTGAAGCATTTGAGATTTACGCCAGCAGTGTGAGGAATCTTTCGGGCTC 561
Qy 444 GAGGAATATGTTGTCAGTGGCGATGAGCTCTCGGCGCACCCAGAGTCTCAGAG 503
Db 562 GAAGATTACTGTGCGCAGTGTGTGGCTGAGCTCAGCGGCACTAGAGAGTCTCGAAG 621
Qy 504 GCCTACATCGCATAGCAGATTGGCAAGAACTTCGAGCAGGAGCGGTGGCCAAAGGAG 563
Db 622 GCATAGTGGCATTCGCTATCTGCGGAAGACATTTTGAGCAGGAAACCTTGGGAAAGAA 681
Qy 564 GTGTCTCTCGAGCAGGAGCATCTGTGCTGCTCGCTCCAGCAGGAGGAGTCTCCTCAGCC 623
Db 682 GTGTCTCTCGAGCAGGAGTGTACTCAGTGCAGCGCCACCTAGAGGATCCCAATGGCT 741
Qy 624 GAGTGGAGTGTCTCGGAAACAGGAGTGTGTGGAGCCCGTCCCTCGAGCCCAATGTATAC 683
Db 742 GAGTGGAGTGTCTAAAGAAATGAAGACATAATTGATCTCTGCGAAGATCGGAATTTTAT 801
Qy 684 ATCAGCGGAGCAGCAGCTGTGTGTGCGAGCGCCGCTTGTGTGACAGCGGCAACTAC 743
Db 802 ATCACTATCGATCAAACTGATCATCAAGCAGGACCACTCTCAGATACAGCGAATTAT 861
Qy 744 ACCTGCTGGGCCAAGAAATCATCTGTGCGCAGCGCTCCGCTGCTGTCTATCTC 803
Db 862 ACCTGTTGTCCAAACATTTGTTCCAGAGGAAAGACCAACAGCCACCGCTCATCTGTG 921
Qy 804 TACGTGAACGTTGGTGTGTGCGAGTGGTGTGCGTGTGAGCGGCGGCGGAGTGTGG 863
Db 922 TATGTTAATGTTGGTGTGTGCGACCTGGCGAGNATGTTGTGTGTAAACAGCGCTGTGG 981
Qy 864 CCGGCTGGCAGAAACGAGCGGAGTGTGACCAACCCGCGCGCTCTCTCAACGCGGCGCT 923
```

Query Match 32.5%; Score 936.8; DB 10; Length 9328;

Best Local Similarity 61.8%; Pred. No. 3.6e-145;

Matches 1631; Conservative 0; Mismatches 937; Indels 72; Gaps 6;

Qy 204 CCGGACCTGCTCCCCACATTCCTGGTGGAGCCCGAGGATGTTACATCGTCAAGAACAAAG 263

Db 322 CCTGAGCCATTACACATTTCTTATTGAGCCCGAGGAGCTTACATTTGAGAGAACAAAG 381

Qy 264 CCAAGTGTCTTGTGTGCAAGGCCGTGCCCGCCAGCAGATCTTCTTCAAGTGCAACGGG 323

Db 382 CTTGTAACCTGTATTGTAAGCCAGCCCTGTCACCCAGATCTTCTTCAAGTGCAACAGC 441

Qy 324 GAGTGGTGGCCAGGTGGACCAAGTGTGAGCGGAGCAGACAGCGGAGCAGTGTGTAG 383

Db 442 GAATGGTTCATCAGAAGGACCAAGTAGTAGAGAGGGGTAGATGAAACCTCCGCTCTC 501

Qy 384 CCGACCATGAGGTCCGCAATTAATGTCTCAAGCAGCAGTCCGAGAGGTGTTTCGGGCTG 443

Db 502 ATTGTGAGAGAGTGAAGCATTTGAGATTTACGCCAGCAGTGTGAGGAATCTTTCGGGCTC 561

Qy 444 GAGGAATATGTTGTCAGTGGCGATGAGCTCTCGGCGCACCCAGAGTCTCAGAG 503

Db 562 GAAGATTACTGTGCGCAGTGTGTGGCTGAGCTCAGCGGCACTAGAGAGTCTCGAAG 621

Qy 504 GCCTACATCGCATAGCAGATTGGCAAGAACTTCGAGCAGGAGCGGTGGCCAAAGGAG 563

Db 622 GCATAGTGGCATTCGCTATCTGCGGAAGACATTTTGAGCAGGAAACCTTGGGAAAGAA 681

Qy 564 GTGTCTCTCGAGCAGGAGCATCTGTGCTGCTCGCTCCAGCAGGAGGAGTCTCCTCAGCC 623

Db 682 GTGTCTCTCGAGCAGGAGTGTACTCAGTGCAGCGCCACCTAGAGGATCCCAATGGCT 741

Qy 624 GAGTGGAGTGTCTCGGAAACAGGAGTGTGTGGAGCCCGTCCCTCGAGCCCAATGTATAC 683

Db 742 GAGTGGAGTGTCTAAAGAAATGAAGACATAATTGATCTCTGCGAAGATCGGAATTTTAT 801

Qy 684 ATCAGCGGAGCAGCAGCTGTGTGTGCGAGCGCCGCTTGTGTGACAGCGGCAACTAC 743

Db 802 ATCACTATCGATCAAACTGATCATCAAGCAGGACCACTCTCAGATACAGCGAATTAT 861

Qy 744 ACCTGCTGGGCCAAGAAATCATCTGTGCGCAGCGCTCCGCTGCTGTCTATCTC 803

Db 862 ACCTGTTGTCCAAACATTTGTTCCAGAGGAAAGACCAACAGCCACCGCTCATCTGTG 921

Qy 804 TACGTGAACGTTGGTGTGTGCGAGTGGTGTGCGTGTGAGCGGCGGCGGAGTGTGG 863

Db 922 TATGTTAATGTTGGTGTGTGCGACCTGGCGAGNATGTTGTGTGTAAACAGCGCTGTGG 981

Qy 864 CCGGCTGGCAGAAACGAGCGGAGTGTGACCAACCCGCGCGCTCTCTCAACGCGGCGCT 923

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2004, 09:36:51 ; Search time 1081 seconds
(without alignments)
11321.978 Million cell updates/sec

Title: US-09-970-944-1

Perfect score: 2881

Sequence: 1 agctggggtccggctgag.....ccttccccacacgggggaga 2881

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_29Jan04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002s.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB ID	Description
1	2881	100.0	2881	6	Abk49422 DNA encod
2	2676.4	92.9	2752	6	Abk37922 cDNA enco
3	2615	90.8	2697	6	Abk52891 Human net
4	2402.6	83.4	3580	6	Abk15169 Human REP
5	2296.2	79.7	2907	4	Aak52261 Human pol
6	2235.8	77.6	3014	2	Aav52940 Rat UNC-5
7	2189.4	76.0	2697	6	Aas16843 Rat netri
8	1570.4	54.5	1787	2	Aav52941 Human UNC
9	1291.2	44.8	9700	4	Aac90958 Plasmid p
10	1193.6	41.4	1321	4	Aah99530 Human pro
11	916	31.8	3646	5	Aas75738 DNA encod
12	889	30.9	2860	6	Abt06279 Human NOV
13	887.4	30.8	2860	6	Abt06280 Human NOV
14	875.8	30.4	2995	6	Abk92062 DNA encod
15	866.4	30.1	2895	6	Abg93898 Human tra
16	853.6	29.6	3884	4	Aas21316 Human cDN
17	853.6	29.6	3884	7	ACd23925 Novel hum
18	853.6	29.6	3884	7	ACA67066 cDNA enco
19	853.6	29.6	3884	7	ACA03675 cDNA enco
20	853.6	29.6	3884	7	Abx89213 DNA encod
21	853.6	29.6	3884	7	ACd41867 Human sec
22	853.6	29.6	3884	7	ACA04096 Human cDN
23	853.6	29.6	3884	8	Ada45664 Novel hum

24	853.6	29.6	3884	8	ADA76095 Human PRO
25	853.6	29.6	3884	8	ADA18745 Human PRO
26	853.6	29.6	3884	8	ADA61368 Homo sapi
27	853.6	29.6	3884	8	ADB19153 Novel hum
28	853.6	29.6	3884	8	ADB27694 cDNA enco
29	853.6	29.6	3884	8	ADA86173 Novel hum
30	853.6	29.6	3884	8	ADB15737 Human PRO
31	853.6	29.6	3884	8	ADA47523 Human PRO
32	853.6	29.6	3884	8	ADA67318 Human PRO
33	853.6	29.6	3884	8	ADB30325 cDNA enco
34	853.6	29.6	3884	8	ADA85621 Novel hum
35	853.6	29.6	3884	8	ADA96833 Human PRO
36	853.6	29.6	3884	8	ADA79137 Human PRO
37	853.6	29.6	3884	8	ADA87276 Novel hum
38	853.6	29.6	3884	8	ADB16478 Human PRO
39	853.6	29.6	3884	8	ADA91570 Human PRO
40	853.6	29.6	3884	8	ADB14633 Novel hum
41	853.6	29.6	3884	8	ADA91570 Human PRO
42	853.6	29.6	3884	8	ADB18594 Novel hum
43	853.6	29.6	3884	8	ADA93809 Human PRO
44	853.6	29.6	3884	8	ADB19705 Novel hum
45	853.6	29.6	3884	8	ADB13017 Human PRO
					ACd98496 Novel hum

ALIGNMENTS

RESULT 1

ABK49422
ID ABK49422 standard; DNA; 2881 BP.
XX
AC ABK49422;
XX
DT 15-JUL-2002 (first entry)
XX
DE DNA encoding human UNC5-like protein NOV1.
XX
KW Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;
KW cell signal processing; metabolic pathway modulation; cancerous tissue;
KW antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;
KW chromosome 13; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 87..2786
FT /*tag= a
FT /product= "Human UNC5-like protein NOV1"
XX

WG200229038-A2.

11-APR-2002.

04-OCT-2001; 2001WO-US031377.

04-OCT-2000; 2000US-0237862P.

(CURA-) CURAGEN CORP.

Herrmann JL, Rastelli L, Shinkets RA;

WPI; 2002-340104/37.

P-PSDB; AAU79939.

Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for treating cardiomyopathy, atherosclerosis, and cancer.

Claim 8; Page 7-8; 180pp; English.

The present invention relates to a new NOVX polypeptide having a 900 (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or 331 (NOV6) residue amino acid sequence, as given in the specification. The novel polypeptide, and its encoding polynucleotide, are used to treat

	CC	cardiomyopathy, atherosclerosis, cancer or a disease related to cell
CC	CC	signal processing and metabolic pathway modulation, in a human. Detecting
CC	CC	the polypeptide or polynucleotide is useful for identifying cancerous
CC	CC	tissue. The antibody can be used to treat diabetes or cancer. The host
CC	CC	cells can be used to produce non-human transgenic animals useful in drug
CC	CC	screening. The present nucleic acid sequence is that of the human UNC5-
CC	CC	like NOV1 gene located on chromosome 13. This sequence encodes the human
CC	CC	UNC5-like protein NOV1 of the invention
XX		
SQ	Sequence 2881 BP; 526 A; 985 C; 868 G; 502 T; 0 U; 0 Other;	
	Query Match 100.0%; Score 2881; DB 6; Length 2881;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 2881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 AGCTGGGGCTCCGGGCTGAGCGCTAAAGCGCCCTCCGCCC CGGGGCCCGGGCGCGG 60	
DB	1 AGCTGGGGCTCCGGGCTGAGCGCTAAAGCGCCCTCCGCCC CGGGGCCCGGGCGCGG 60	
QY	61 GCCCGCCGCTGCGCCCGCGGCCATGGCGCTCGGCGCCGGCCCTGTGGCCACGCTCC 120	
DB	61 GCCCGCCGCTGCGCCCGCGGCCATGGCGCTCGGCGCCGGCCCTGTGGCCACGCTCC 120	
QY	121 TGGCATAGTCCTCGCCCTTGGGCTCCGCGGCTCGGGTGCCACGAGTCCACCGTGG 180	
DB	121 TGGCATAGTCCTCGCCCTTGGGCTCCGCGGCTCGGGTGCCACGAGTCCACCGTGG 180	
QY	181 CCAACCCAGTCCTGGTGCCAAACCCGGAACCTGCTTCCCACTTCTTGTTGAGCCCGGAGG 240	
DB	181 CCAACCCAGTCCTGGTGCCAAACCCGGAACCTGCTTCCCACTTCTTGTTGAGCCCGGAGG 240	
QY	241 ATGTGTACATCGTCAAGAACAGCCAGTGTCTTGTGTCAAGGCCCTGCGCCGACGCG 300	
DB	241 ATGTGTACATCGTCAAGAACAGCCAGTGTCTTGTGTCAAGGCCCTGCGCCGACGCG 300	
QY	301 AGATCTTCTTCAAGTGCAAACCGGGGAGTGGTGCGCCAGGTGGACCACTGTATCGAGCGCA 360	
DB	301 AGATCTTCTTCAAGTGCAAACCGGGGAGTGGTGCGCCAGGTGGACCACTGTATCGAGCGCA 360	
QY	361 GCACAGACGGGAGCAGTGGTGGAGCCGACATGGAGGTCGGCATTAATCTCTCAAGGACG 420	
DB	361 GCACAGACGGGAGCAGTGGTGGAGCCGACATGGAGGTCGGCATTAATCTCTCAAGGACG 420	
QY	421 AGTGCAGAAAGTGTTCGGGCTTGAGGAAATCTGGTGCCAGTGGCGATGGAGCTCTCT 480	
DB	421 AGTGCAGAAAGTGTTCGGGCTTGAGGAAATCTGGTGCCAGTGGCGATGGAGCTCTCT 480	
QY	481 CGGGACCAACCAAGAGTCAAGAGGCTCATCCGCATAGCCAGATGGCCAGAGACTTCG 540	
DB	481 CGGGACCAACCAAGAGTCAAGAGGCTCATCCGCATAGCCAGATGGCCAGAGACTTCG 540	
QY	541 AGCAGGACCGCTGGCCAGGAGGTGTCCTCGAGCAGGGCATCTGTCTGCCCTGCGGTC 600	
DB	541 AGCAGGACCGCTGGCCAGGAGGTGTCCTCGAGCAGGGCATCTGTCTGCCCTGCGGTC 600	
QY	601 CACCGGAGGGCATCCCTCCAGCCGAGTGGAGTGGCTCCGGAACGAGGACCTTGTGTGAC 660	
DB	601 CACCGGAGGGCATCCCTCCAGCCGAGTGGAGTGGCTCCGGAACGAGGACCTTGTGTGAC 660	
QY	661 CGTCCCTGACCCCAATGTATACATCAGCGGGAGCACGCTGGTGGTGGCCAGGCGCC 720	
DB	661 CGTCCCTGACCCCAATGTATACATCAGCGGGAGCACGCTGGTGGTGGCCAGGCGCC 720	
QY	721 GCCTTGTGCACACGGCCAACTACACCTGGTGGCCAAAGAACACTCGTGGCACGTCGCCGA 780	
DB	721 GCCTTGTGCACACGGCCAACTACACCTGGTGGCCAAAGAACACTCGTGGCACGTCGCCGA 780	
QY	781 GGGCTTCCGCTGTCATCGTCTACGTGNAACGGTGGGTGGTTCAGCTGGACCGAGTGGT 840	
DB	781 GGGCTTCCGCTGTCATCGTCTACGTGNAACGGTGGGTGGTTCAGCTGGACCGAGTGGT 840	
QY	841 CGTCTTGCAGCGCCAGCTGTGGCGCGGGCTGGCAGAAACGGAGCCGGAGTGCACCAAC 900	

QY 2139 CTGAGAGCAGCTGGGGGACAGCTGATCCAGAGGCCACCGGTCCTGCACTTCAAGGAC 2198
 Db 2095 CTGAGAGCAGCTGGGGGACAGCTGATCCAGAGGCCACCGGTCCTGCACTTCAAGGAC 2154
 QY 2199 AGTTACACACACCTGGGCTATCATCCACAGTGTGCCAGCTCCCTGTGGAAGAGTAAG 2258
 Db 2155 AGTTACACACACCTGGGCTATCATCCACAGTGTGCCAGCTCCCTGTGGAAGAGTAAG 2214
 QY 2259 CTCCTGTGACACTACAGGAGATCCCTTTTATCAGATCTGGAATGGCAGCGGTAC 2318
 Db 2215 CTCCTGTGACACTACAGGAGATCCCTTTTATCAGATCTGGAATGGCAGCGGTAC 2274
 QY 2319 TTGCACCTGCACCTTCAACCTGGAGCGTGTGAGCCCGCAGCACTAGTACCTGGCTGCAAG 2378
 Db 2275 TTGCACCTGCACCTTCAACCTGGAGCGTGTGAGCCCGCAGCACTAGTACCTGGCTGCAAG 2334
 QY 2379 CTGTGGGTGTGGAGGTGGAGGGGAGCGGCGAGCTTCAGATCAACTTCAACATCACC 2438
 Db 2335 CTGTGGGTGTGGAGGTGGAGGGGAGCGGCGAGCTTCAGATCAACTTCAACATCACC 2394
 QY 2439 AAGGACACAAAGTTTGTGAGCTGTGCTGTGAGAGTGAAGCGGGTCCAGCCCTG 2498
 Db 2395 AAGGACACAAAGTTTGTGAGCTGTGCTGTGAGAGTGAAGCGGGTCCAGCCCTG 2454
 QY 2499 GTGGGCCCGAGTGCCTTCAAGATCCCTTCTCATTCGCGAGAGATAATTTCCAGCCTG 2558
 Db 2455 GTGGGCCCGAGTGCCTTCAAGATCCCTTCTCATTCGCGAGAGATAATTTCCAGCCTG 2514
 QY 2559 GACCCACCTGTAGGGGGTGGCTGCTGTGAGAGTGAAGCGGGTCCAGCCCTG 2618
 Db 2515 GACCCACCTGTAGGGGGTGGCTGCTGTGAGAGTGAAGCGGGTCCAGCCCTG 2574
 QY 2619 AGCCATCTCAGCTTCTTTGCTTCCAGCCCGAGCCGAGCCATGATCTTCAACCTGTGG 2678
 Db 2575 AGCCATCTCAGCTTCTTTGCTTCCAGCCCGAGCCGAGCCATGATCTTCAACCTGTGG 2634
 QY 2679 GAGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGAGAGTGGTGGACTG 2738
 Db 2635 GAGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGAGAGTGGTGGACTG 2694
 QY 2739 GGCAGCCAGACCTGGCTTCTTCAAGTGTGGAGGCTGAGTGTGAGCGCGGCCAG 2796
 Db 2695 GGCAGCCAGACCTGGCTTCTTCAAGTGTGGAGGCTGAGTGTGAGCGCGGCCAG 2752

RESULT 3

ID ABK52891 standard; DNA; 2697 BP.
 AC ABK52891;
 DT 27-AUG-2002 (first entry)
 DE Human netrin binding membrane receptor UNC5H-1 DNA sequence #1.
 KW Netrin binding membrane receptor; receptor; UNC5H-1; gene; ds; human;
 KW neotrophic; neuroprotective; cytoskeletal; antiparkinsonian;
 KW cerebroprotective; cancer; central nervous system; CNS; stroke;
 KW Parkinson's disease; multiple sclerosis; Alzheimer's disease.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT 1..2697
 FT /*tag= a
 FT /product= "Netrin binding membrane receptor UNC5H-1"
 XX
 PN W0200233080-A2.
 XX
 PD 25-APR-2002.
 XX
 PF 15-OCT-2001; 2001WO-EP011891.

XX PR 16-OCT-2000; 2000US-0240061P.
 PA (FARB) BAYER AG.
 XX Koehler RH;
 XX WPI; 2002-463314/49.
 DR P-ESDB; AAU97899.
 XX
 PT Novel human netrin binding membrane receptor polypeptide and
 PT polynucleotides for identifying modulating agents useful in treating
 PT diseases e.g Parkinson's disease, multiple sclerosis, stroke,
 PT Alzheimer's disease.
 XX
 PS Claim 1; Fig 1; 94pp; English.
 XX
 CC This invention relates to the DNA and protein sequences of a novel
 CC purified human netrin binding membrane receptor, UNC5H-1. The DNA
 CC sequence of the invention is useful as a probe for detecting a nucleic
 CC acid encoding the UNC5H-1 protein in a biological sample. The sequences
 CC of the invention are useful to screen for agents which decrease the
 CC activity of the UNC5H-1 protein. The sequences are also useful for
 CC screening agents which regulate (modulate) the activity of the protein of
 CC the invention. A pharmaceutical composition containing the protein of the
 CC invention or a reagent that modulates the activity of the protein of the
 CC may be useful for treating a UNC5H-1 dysfunction related disease such as
 CC cancer or a central nervous system (CNS) disorders (e.g. Parkinson's
 CC disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion
 CC proteins comprising the UNC5H-1 protein and Alzheimer's disease
 CC antibodies and for in various assay systems, and the protein can be used
 CC as a bait protein in a two-hybrid assay or three-hybrid assay. The method
 CC of the invention is useful for detecting a coding sequence for the UNC5H-
 CC 1 protein. The present sequence represents a DNA sequence encoding the
 CC human netrin binding membrane receptor UNC5H-1 protein of the invention
 XX

Sequence 2697 BP; 503 A; 906 C; 807 G; 481 T; 0 U; 0 Other;

Query Match 90.8%; Score 2615; DB 6; Length 2697;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 2669; Conservative 0; Mismatches 25; Indels 9; Gaps 3;
 QY 87 ATGGCGCTCGGCGCGGCTGTGGCCAGCGCTCTCTGGGCATAGTCTCGCGCTTGGCTC 146
 Db 1 ATGGCGCTCGGCGCGGCTGTGGCCAGCGCTCTCTGGGCATAGTCTCTCGCGCTTGGCTC 60
 QY 147 CGCGGCTCGGCTCGGCGCGGCTGTGGCCAGCGCTCTCTGGGCATAGTCTCTCGCGCTTGGCTC 206
 Db 61 CGCGGCTCGGCTCGGCGCGGCTGTGGCCAGCGCTCTCTGGGCATAGTCTCTCGCGCTTGGCTC 120
 QY 207 GACCTGCTTCCCACTTCTCTGGTGGAGCCCGAGGATGTGTACATGCTCAAGAACAGGCCA 266
 Db 121 GACCTGCTTCCCACTTCTCTGGTGGAGCCCGAGGATGTGTACATGCTCAAGAACAGGCCA 180
 QY 267 GTGCTGCTGTGTGCAAGCCCTGCGCCAGCGAGATCTTCTCAAGTGAACGGGAG 326
 Db 181 GTGCTGCTGTGTGCAAGCCCTGCGCCAGCGAGATCTTCTCAAGTGAACGGGAG 240
 QY 327 TGGGTCGCGCAGGTGGACCACTGATCGAGCGCAGCAGACGGGAGGAGTGTGAGCCG 386
 Db 241 TGGGTCGCGCAGGTGGACCACTGATCGAGCGCAGCAGACGGGAGGAGTGTGAGCCG 300
 QY 387 ACCATGGAGGTCCGCAATTAATGCTCAAGGCGACGAGGTCTCGGAGGTGTTCGGGTGGAG 446
 Db 301 ACCATGGAGGTCCGCAATTAATGCTCAAGGCGACGAGGTCTCGGAGGTGTTCGGGTGGAG 360
 QY 447 GAATACTGGTCCAGTGGTGCATGAGCTCTCTCGGACCAACCAAGAGTCAAGAGGCC 506
 Db 361 GAATACTGGTCCAGTGGTGCATGAGCTCTCTCGGACCAACCAAGAGTCAAGAGGCC 420
 QY 507 TACATCCGCATAGCCAGATTGGCGAAGAACTTCGAGCAGGAGGCCCTGGCCAGAGGTG 566
 Db 421 TACATCCGCATAGCCCTATTTCGCAAGAACTTCGAGCAGGAGGCCCTGGCCAGAGGTG 480

QY	567	TCCCTGGAGCAGGGCATCGTGTCCCTCGCTCCACCGAGGGCATCCCTCCAGCCGAG	626
Db	481	TCCCTTGGAGCAGGGCATCGTGTCCCTCGCTCCACCGAGGGCATCCCTCCAGCCGAG	540
QY	627	GTGAGTGGCTCCGGAACGAGGACCTGTGTGGACCGTCCCTGGACCCCAATGTATACATC	686
Db	541	GTGAGTGGCTCCGGAACGAGGACCTGTGTGGACCGTCCCTGGACCCCAATGTATACATC	600
QY	687	ACGGGGAGCACAGCTGTGTGTGGACAGGCGCGCTTGTGTGACAGGCGCAACTACACC	746
Db	601	ACGGGGAGCACAGCTGTGTGTGGACAGGCGCGCTTGTGTGACAGGCGCAACTACACC	660
QY	747	TGCGTGCCCAAGAACATCTGTGGCAGCTGCGCAGCGCCCTCCGCTGCTCATCGTCTAC	806
Db	661	TGCGTGCCCAAGAACATCTGTGGCAGCTGCGCAGCGCCCTCCGCTGCTCATCGTCTAC	720
QY	807	GTGAAACGGTGGGTGTGACGTGGAACGAGTGTCTCGTCTGTGACGCGCAGCTGTGGCGC	866
Db	721	GTGAAACGGTGGGTGTGACGTGGAACGAGTGTCTCGTCTGTGACGCGCAGCTGTGGCGC	780
QY	867	GGCTGGCAGAAAACGAGCCGAGCTGCACCAACCCGCGCCTCTCAACGGGGCGCTTTC	926
Db	781	GGCTGGCAGAAAACGAGCCGAGCTGCACCAACCCGCGCCTCTCAACGGGGCGCTTTC	840
QY	927	TGTGAGGGGCGAATGTCCATGACCGCACCGTCTCCTCTCTGTCTGTGTGTGACGCGC	986
Db	841	TGTGAGGGGCGAATGTCCAGAA---AACAGCCTGGCGCACCCCTGTGTGCCAGTGGACGCGC	897
QY	987	AGCTGGAGCCGCTGGAGCAAGTGTGTGGCTGTGGCTGTGACTGCACCCACTGGCGGAGC	1046
Db	898	AGCTGGAGCCGCTGGAGCAAGTGTGTGGCTGTGGCTGTGACTGCACCCACTGGCGGAGC	957
QY	1047	CGTGAGTGTCTGTGACCCAGCACCCCGCAACCGAGGGGAGGAGTGCACGAGGCACTGACCTG	1106
Db	958	CGTGAGTGTCTGTGACCCAGCACCCCGCAACCGAGGGGAGGAGTGCACGAGGCACTGACCTG	1017
QY	1107	GACACCCGCAACTGTACAGTGAACCTGTGTGTACAGTGTCTTGTGGCCCTGAGGACGTG	1166
Db	1018	GACACCCGCAACTGTACAGTGAACCTGTGTGTGTACAGTGTCTTGTGGCCCTGAGGACGTG	1077
QY	1167	GCCCTCATGTGGGCTCATCGCGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTC	1226
Db	1078	GCCCTCATGTGGGCTCATCGCGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTC	1137
QY	1227	ATCCTCGTTTATTCGGGAAGAGAGGGCTGTGACTCAGATGTGGCTGACTCGTCCATT	1286
Db	1138	ATCCTCGTTTATTCGGGAAGAGAGGGCTGTGACTCAGATGTGGCTGACTCGTCCATT	1197
QY	1287	CTCACCTCAGCTTCCAGCCCGTCAAGTCAAGCCGAGCAAGCAGACACCCCACTCTG	1346
Db	1198	CTCACCTCAGCTTCCAGCCCGTCAAGTCAAGCCGAGCAAGCAGACACCCCACTCTG	1257
QY	1347	CTCACCTCAGCGGCACTCAG---CACCAACCACTTACCAGGGCAGTCTCTGTCCC	1403
Db	1258	CTCACCTCAGCGGCACTCAGCAACCACTTACCAGGGCAGTCTCTGTCCC	1317
QY	1404	CGCAGGATGGCCCGACCCCAAGTTCCAGTGTACCAATGAGGACCTGTGCTCAGCCCTG	1463
Db	1318	CGCAGGATGGCCCGACCCCAAGTTCCAGTGTACCAATGAGGACCTGTGCTCAGCCCTG	1377
QY	1464	GCTGCGCGCGCACACACTGCACACAGCTCTCCCACTCTGAGGCGCAGGAGTTCGTC	1523
Db	1378	GCTGCGCGCGCACACACTGCACACAGCTCTCCCACTCTGAGGCGCAGGAGTTCGTC	1437
QY	1524	TCCCGCTCTCCACCGAATCTACTTTCGGCTCCCTGTGCTCCCGCAGGCAACAGCATGACC	1583
Db	1438	TCCCGCTCTCCACCGAATCTACTTTCGGCTCCCTGTGCTCCCGCAGGCAACAGCATGACC	1497
QY	1584	TATGGACCTTCAACTTCTCGGGGCGGCTGTGTATGCCCTAATACAGGTATCAGGCTC	1643
Db	1498	TATGGACCTTCAACTTCTCGGGGCGGCTGTGTATGCCCTAATACAGGTATCAGGCTC	1557

1644	QY	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACTCAGCTGCAC	1703
1558	Db	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACTCAGCTGCAC	1617
1704	QY	AAGCGGAAGAGTGAAGTTGCCCTAGCTGAGTGCATGACAGCCCTGCTGAGTCCCATCGTT	1763
1618	Db		1677
1764	QY	AGCTGTGGACCCCTGGCGTCTTGCTCACCCGGCCAGTGCATCTGGCTATGACCACTGT	1823
1678	Db	AGCTGTGGACCCCTGGCGTCTTGCTCACCCGGCCAGTGCATCTGGCTATGACCACTGT	1737
1824	QY	GGGAGCCAGCCCTGACAGCTGGAGCCTGGCGCTCAAAAAGAGTGCCTGGAGGGCAGC	1883
1738	Db	GGGAGCCAGCCCTGACAGCTGGAGCCTGGCGCTCAAAAAGAGTGCCTGGAGGGCAGC	1797
1884	QY	TGGAGCAGAGTGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAG	1943
1798	Db		1854
1944	QY	CTGGAGCCAGTGCCTACTACGTCTTACCAGACAGCTGGGCGCGTTGCCCTGTGGGA	2003
1855	Db	CTGGAGCCAGTGCCTACTACGTCTTACCAGACAGCTGGGCGCGTTGCCCTGTGGGA	1914
2004	QY	GAGGCCCTAGCTGGTGGCGGCAAGCGGCTCAAGCTGCTTCTGTTGGCGCGGTGGCC	2063
1915	Db	GAGGCCCTAGCGTGGCTGGCGCAAGCGGCTCAAGCTGCTTCTGTTGGCGCGGTGGCC	1974
2064	QY	TGCACCTCCCTCAGGTACACATCCGGGTCTACTGCTGCTGACGACACCCACGATGACATC	2123
1975	Db		2034
2124	QY	AAGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGGAAGCTGATCCAGGAGCCACGGTTC	2183
2035	Db	AAGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGGAAGCTGATCCAGGAGCCACGGTTC	2094
2184	QY	CTGCATTCGAAGACAGTTACCAACACTTGGCGCTATCCATCCACGATGTGCCACGCTCC	2243
2095	Db	CTGCATTCGAAGACAGTTACCAACACTTGGCGCTATCCATCCACGATGTGCCACGCTCC	2154
2244	QY	CTGTGGAAGAGTAAAGCTCCTTGTGACCTACAGAGATCCCTTTTATCACATCTGGAAT	2303
2155	Db	CTGTGGAAGAGTAAAGCTCCTTGTGACCTACAGAGATCCCTTTTATCACATCTGGAAT	2214
2304	QY	GGCAGCAGCGGTACTTTGCACTGCACCTTCAACCTTGAGCGTGTGACGCCACGACTAGT	2363
2215	Db	GGCAGCAGCGGTACTTTGCACTGCACCTTCAACCTTGAGCGTGTGACGCCACGACTAGT	2274
2364	QY	GACCTGCGCTGAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGAGAGCTTCAGCATC	2423
2275	Db	GACCTGCGCTGAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGAGAGCTTCAGCATC	2334
2424	QY	AACCTCAACATCACCAGAGCACAAAGTTTGTGAGCTGTGCTTGGAGAGTGAAGCG	2483
2335	Db	AACCTCAACATCACCAGAGCACAAAGTTTGTGAGCTGTGCTTGGAGAGTGAAGCG	2394
2484	QY	GGGTGCCAGCCCTGTGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAAG	2543
2395	Db	GGGTGCCAGCCCTGTGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAAG	2454
2544	QY	ATAATTTCCAGCTGGAGCCACCTGTGAGGGGGGTGCCAGTGGCGGACTCTGGGCCAG	2603
2455	Db	ATAATTTCCAGCTGGAGCCACCTGTGAGGGGGGTGCCAGTGGCGGACTCTGGGCCAG	2514
2604	QY	AAACTCCACCTGGACAGCATCTCAGCTTCTTTTGCCTCAAGCGCCAGCCCAAGCCATG	2663
2515	Db	AAACTCCACCTGGACAGCATCTCAGCTTCTTTTGCCTCAAGCGCCAGCCCAAGCCATG	2574
2664	QY	ATCCTCAACTGTGGGAGGCGCGCACTTCCCAAGCGCAACTCAGCAGCTGGCTGCA	2723
2575	Db	ATCCTCAACTGTGGGAGGCGCGCACTTCCCAAGCGCAACTCAGCAGCTGGCTGCA	2634
2724	QY	GCAGTGGCTGCACTGGGCCAGCCAGACGGTGGCTCTTCAAGTGTGGAGCGCTAGTGC	2783

Db 2635 CGAGTGGCTGACTGGCCAGCAGACGCTGGCTCTTACAGTGTGGAGGCTGAGTGC 2694
 Qy 2784 TGA 2786
 Db 2695 TGA 2697
 RESULT 4
 ABK15169
 ID ABK15169 standard; DNA; 3580 BP.
 AC ABK15169;
 XX 23-APR-2002 (first entry)
 DT
 XX Human REPTR 1 cDNA sequence.
 DE
 XX REPTR; human; antiinflammatory; cytostatic; immunosuppressive; antiviral;
 KW anti-HIV; antiarthritic; anticonvulsant; nootropic; neuroprotective;
 KW antiallergic; antibody; immunogen; endometriosis;
 KW gastrointestinal disorder; gastritis; oesophageal carcinoma;
 KW Crohn's disease; irritable bowel syndrome; ulcerative colitis;
 KW endocrine disorder; hypothyroidism; Kallman's disease;
 KW autoimmune disease; inflammatory disease; infertility; receptor;
 KW acquired immune deficiency syndrome; AIDS; rheumatoid arthritis; allergy;
 KW osteoarthritis; diabetes mellitus; multiple sclerosis;
 KW systemic lupus erythematosus; cell proliferative disorder; cancer;
 KW developmental disorder; Duchenne muscular dystrophy; gene;
 KW Becker muscular dystrophy; neurological disorder; epilepsy;
 KW Alzheimer's disease; Huntington's disease; reproductive disorder; ss.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT 4.2532
 FT /*tag= a
 FT /*product= "REPTR1 protein"
 XX
 PN WO200198354-A2.
 XX
 XX 27-DEC-2001.
 XX
 XX 21-JUN-2001; 2001WO-US019942.
 XX
 XX 21-JUN-2000; 2000US-0214027P.
 PR 25-AUG-2000; 2000US-0228045P.
 PR 12-DEC-2000; 2000US-0255104P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Griffin JA, Kallick DA, Tribouley CM, Yue H, Nguyen DB, Tang YT,
 PI Lal P, Policky JL, Azinai Y, Lu DM, Gaul R, Yao MG, Burford N,
 PI Hafalia AJA, Baughn MR, Bandman O, Patterson C, Yang J, Xu Y,
 PI Gandhi AR, Warren BA, Ding L, Sanjanwala MS, Duggan BM, Lu Y;
 XX
 DR WPI; 2002-090432/12.
 DR P-PSDB; AAU17818.
 XX
 XX Twelve human receptors (referred to as REPTR-1 to REPTR-12), useful in
 PT the diagnosis, treatment and prevention of gastrointestinal (e.g.
 PT gastritis), autoimmune/inflammatory (e.g. osteoarthritis) and cell
 PT proliferative (e.g. cancer) disorders.
 XX
 PS Claim 57; Page 142-143; 157pp; English.
 XX
 CC This invention relates to twelve human receptors cDNA sequences referred
 CC to as REPTR-1 to REPTR-12), and the proteins encoded thereby. The
 CC proteins of the invention may have antiinflammatory, cytostatic,
 CC immunosuppressive, antiviral, anti-HIV, antiarthritic, muscular active
 CC general, anticonvulsant, nootropic, neuroprotective, antiallergic
 CC activities. The sequences of the invention may be used to produce REPTR
 CC agonists or antagonists, and the protein sequences may be used to raise

CC anti-REPTR antibodies. These molecules and the REPTR polynucleotides and
 CC polypeptides of the invention are useful in the diagnosis, treatment and
 CC prevention of gastrointestinal (e.g. gastritis, oesophageal carcinoma,
 CC Crohn's disease, irritable bowel syndrome, ulcerative colitis), endocrine
 CC (e.g. hypothyroidism, Kallman's disease), autoimmune/inflammatory
 CC (e.g. acquired immune deficiency syndrome (AIDS), rheumatoid arthritis,
 CC allergies, osteoarthritis, diabetes mellitus, multiple sclerosis,
 CC systemic lupus erythematosus), cell proliferative (e.g. cancer),
 CC developmental (e.g. Duchenne and Becker muscular dystrophy), neurological
 CC (e.g. epilepsy, Alzheimer's disease, Huntington's disease) and
 CC reproductive (e.g. infertility, endometriosis) disorders. Numerous other
 CC examples of each disorder are given in the specification. The present
 CC sequence represents the human REPTR1 cDNA sequence of the invention
 XX
 SQ Sequence 3580 BP; 670 A; 1233 C; 1025 G; 652 T; 0 U; 0 Other;
 Query Match 83.4%; Score 2402.6; DB 6; Length 3580;
 Best Local Similarity 93.4%; Pred. No. 0;
 Matches 2615; Conservative 0; Mismatches 9; Indels 177; Gaps 3;
 Qy 84 GCATGCGCCGTCGGGCGGCGCTGTGGCCAGCGCTCTGGGCATAGTCTCGCGCTGG 143
 Db 1 GCATGCGCCGTCGGGCGGCGCTGTGGCCAGCGCTCTGGGCATAGTCTCGCGCTGG 60
 Qy 144 CTCGCGGCTCGGGTGCAGAGAGTGCCACCGTGGCCAAACCGCTGTGGTGGCCAA 203
 Db 61 CTCGCGGCTCGGGTGCAGAGAGTGCCACCGTGGCCAAACCGCTGTGGTGGCCAA 120
 Qy 204 CCGGACCTGCTCCCACTTCTGTTGGAGCGGAGTGATGTACATCGTCAAGAACAG 263
 Db 121 CCGGACCTGCTCCCACTTCTGTTGGAGCGGAGTGATGTACATCGTCAAGAACAG 180
 Qy 264 CAGTGTCTGTTGTGCAAGGCGCTGCCCGCCAGCAGATCTTCTCAAGTCAACGGG 323
 Db 181 CAGTGTCTGTTGTGCAAGGCGCTGCCCGCCAGCAGATCTTCTCAAGTCAACGGG 240
 Qy 324 GAGTGGTGGCCAGGTGGACACAGTGTGCGGCGAGCAGACAGCGGAGCAGTGTGAG 383
 Db 241 GAGTGGTGGCCAGGTGGACACAGTGTGCGGCGAGCAGACAGCGGAGCAGTGTGAG 300
 Qy 384 CCGACCATGGAGTCCGCATTATATGTCATCAAGCAGCAGTGTGCGGAGTGTGGGCTG 443
 Db 301 CCGACCATGGAGTCCGCATTATATGTCATCAAGCAGCAGTGTGCGGAGTGTGGGCTG 360
 Qy 444 GAGGAATCTGTTGCGCAGTGTGCGGCGAGTGTGCGGCGAGCAGACAGCGGAGTGTGAG 503
 Db 361 GAGGAATCTGTTGCGCAGTGTGCGGCGAGTGTGCGGCGAGCAGACAGCGGAGTGTGAG 420
 Qy 504 GCCTACATCCGATACCGAGATTGCGCAAGACTTTCGAGCAGGAGCCGCTGGCCAAAGGAG 563
 Db 421 GCCTACATCCGATACCGAGATTGCGCAAGACTTTCGAGCAGGAGCCGCTGGCCAAAGGAG 480
 Qy 564 GTGTCTCTGAGCAGGAGCAGTGTGCGGCGAGTGTGCGGCGAGCAGACAGCGGAGTGTGAG 623
 Db 481 GTGTCTCTGAGCAGGAGCAGTGTGCGGCGAGTGTGCGGCGAGCAGACAGCGGAGTGTGAG 540
 Qy 624 GAGTGGAGTGTGCTCGGGAACGAGGAGCCTGTGCGGAGCCCGCTCGGAGCCCAATGATAC 683
 Db 541 GAGTGGAGTGTGCTCGGGAACGAGGAGCCTGTGCGGAGCCCGCTCGGAGCCCAATGATAC 600
 Qy 684 ATCAGCGGGAGCAGCGCTGTGTTGGAGCAGCGGCGCTTGTGACACGGCCCAACTAC 743
 Db 601 ATCAGCGGGAGCAGCGCTGTGTTGGAGCAGCGGCGCTTGTGACACGGCCCAACTAC 660
 Qy 744 ACCTGCTGGCCCAAGAACATCGTGGCAGTGTGCGGAGCCCGCTCGGAGCCCAATGATAC 803
 Db 661 ACCTGCTGGCCCAAGAACATCGTGGCAGTGTGCGGAGCCCGCTCGGAGCCCAATGATAC 720
 Qy 804 TACGTGAACGTTGGTGTGTCGACGTGGACCGAGTGTGCGGAGCCCGCTCGGAGCCCAACTAC 863
 Db 721 TAC----- 723
 Qy 864 CCGGCTGGCAGAACCGGAGCGGAGTGTGACCAACCGCGGCGCTCTCAACGGGCGCGT 923

Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorder; arthritis; inflammation; ss.

Homo sapiens.

WO200157190-A2.

09-AUG-2001.

05-FEB-2001; 2001WO-US004098.

03-FEB-2000; 2000US-00496914.

27-APR-2000; 2000US-00560875.

20-JUN-2000; 2000US-00598075.

19-JUL-2000; 2000US-00620325.

01-SEP-2000; 2000US-00654936.

15-SEP-2000; 2000US-00863561.

20-OCT-2000; 2000US-00693325.

30-NOV-2000; 2000US-00728422.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

Xue AU, Yang Y, Wejhrman T, Goodrich R;

WPI; 2001-476283/51.

P-PSDB; AAM79128.

Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy.

Claim 1; Page 2691-2694; 6221pp; English.

The invention relates to polynucleotides (AAM78323-AAM80302) that exhibit activity relating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AAM78323), 2111 (AAM78324) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication

Sequence 2907 BP; 552 A; 966 C; 881 G; 508 T; 0 U; 0 Other;

Query Match

Best Local Similarity 79.7%; Score 2296.2; DB 4; Length 2907;

Matches 2565; Conservative 0; Mismatches 38; Indels 186; Gaps 6;

QY	133	TCGCGCTTGGTCCCGGCTCGGGTCCAGCAGAGTGCACCGGCGCCACCCAGTGC	192
Db	170	TGGCGGCGACAGCAGCAGTGGAGCGCCAGCAGAGTGCCCGTGGCCACCCAGTGC	229
QY	193	CTGGTGCACACCGGACCTGCTTCCCACTTCTGGTGGAGCCCGAGGATGTATCATCG	252
Db	230	CTGGTGCACACCGGACCTGCTTCCCACTTCTGGTGGAGCCCGAGGATGTATCATCG	289
QY	253	TCAGAACACAGCAGT	312
Db	290	TCAGAACACAGCAGT	349
QY	313	AGTGAACCGGGAGTGGGTGGCGCAGGTGGACACCGTGTATCGAGCGCAGCAGCGGA	372
Db	350	AGTGAACCGGGAGTGGGTGGCGCAGGTGGACACCGTGTATCGAGCGCAGCAGCGGA	409
QY	373	GCAGTGGTGAGCGGACCATGGAGGTCCGCAATTAATGTCTCAAGGAGGAGGTGAGAGG	432

Db	410	GCAGTGGGCTGCCACCATGAGGTCCCATTAATGTCTCAAGGAGGAGGTCCGAGAGG	469
QY	433	TGTTGGGCTGGAGGAATACTGGTGGCATGGTGGCATGGTGGCATGGTGGCATGGTGG	492
Db	470	TGTTGGGCTGGAGGAATACTGGTGGCATGGTGGCATGGTGGCATGGTGGCATGGTGG	529
QY	493	AGAGTCAGAGGCTTACATCCGATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCGCC	552
Db	530	AGAGTCAGAGGCTTACATCCGATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCGCC	589
QY	553	TGGCAAGAGGTCTCCCTGGAGCAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG	612
Db	590	TGGCAAGAGGTCTCCCTGGAGCAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG	649
QY	613	TCCCTCCAGCGGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	672
Db	650	TCCCTCCAGCGGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	709
QY	673	CCAATGTATACATCACCGGAGCAGCCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	732
Db	710	CCAATGTATACATCACCGGAGCAGCCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	769
QY	733	CGGCCAATACCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	792
Db	770	CGGCCAATACCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	829
QY	793	CTGTATCTGTCTAC	806
Db	830	CTGTATCTGTCTAC	889
QY	807	-----GTGAACGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	825
Db	890	CCCTGGGACGTGACATGTGGTGTCTCTGTCTGGCGCCAGTGAACGGTGGTGGTGGTGG	949
QY	826	CGTGACCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	885
Db	950	CGTGACCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1009
QY	886	GGAGCTGCACCAACCGGCGCTCTCAACGGGGCGCTTCTGTGTGGGGGCGCAATGTCC	945
Db	1010	GGAGCTGCACCAACCGGCGCTCTCAACGGGGCGCTTCTGTGTGGGGGCGCAATGTCC	1069
QY	946	ATGACCGACCGTCTCTCTGT	1005
Db	1070	AGAA---AACAGCGTGGCGCACCTGTGTGGCGCAGTGGACGCGCAGTGGACCGCGTGGAG	1126
QY	1006	AGTGTGGCGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1065
Db	1127	AGTGTGGCGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1186
QY	1066	CACCCGCAACCGGAGGAGTGGCAGGCGACTGACCTGGACACCGCAACTGTGTACCA	1125
Db	1187	CACCCGCAACCGGAGGAGTGGCAGGCGACTGACCTGGACACCGCAACTGTGTACCA	1246
QY	1126	GTGACCTCTGTGTACACA	1143
Db	1247	GTGACCTCTGTGTACACA	1306
QY	1144	---GTGCTTCTGGCGTGGAGCGTGGCGCTTATGTGGCGCTCATGGCGTGGCGTCT	1200
Db	1307	CAGTGTCTTCTGGCGTGGAGCGTGGCGCTTATGTGGCGCTCATGGCGTGGCGTCT	1366
QY	1201	GCCTGGTCTGT	1260
Db	1367	GCCTGGTCTGT	1426
QY	1261	ACTCAGATGTGGTGAATCTGT	1320
Db	1427	ACTCAGATGTGGTGAATCTGT	1486
QY	1321	CCAGCAAGGACAGACACCCCATCTGTCTACCATCCAGCGGACCTCAG---CACACCA	1377

1487	DB	CCAGCAAAAGCAGACAACCCCAATCTGCTCAACATCAAGCCGGACCTCAGACCAACCA	1546
1378	QY	CCACCTTACCAGGCGAGTCTCTGTCCCGGCAGGATGGGCCCAAGCTTCAGAGTCA	1437
1547	DB	CCACCTTACCAGGCGAGTCTCTGTCCCGGCAGGATGGGCCCAAGCTTCAGAGTCA	1606
1438	QY	CCAATGGGCACTGTCTCAGCCCCCTGGGTGGGGCGGCACACATGCACACAGCTCTC	1497
1607	DB	CCAATGGGCACTGTCTCAGCCCCCTGGGTGGGGCGGCACACATGCACACAGCTCTC	1666
1498	QY	CCACCTCTGAGGCCGAGGAGTTGGTCTCCCGCTCTCCACCCAGAACTCTCCGCTCC	1557
1667	DB	CCACCTCTGAGGCCGAGGAGTTGGTCTCCCGCTCTCCACCCAGAACTCTCCGCTCC	1726
1558	QY	TGCCCCGAGGCACACGAACTAGCACTATGGGACCTTCAACTCTCGGGGGCCGGCTGA	1617
1727	DB	TGCCCCGAGGCACACGAACTAGCACTATGGGACCTTCAACTCTCGGGGGCCGGCTGA	1786
1618	QY	TGATCCCTTAATCAGGTATCAGCTCTCTCATCCCCCAGATGCCTATACCCCGAGGGA	1677
1787	DB	TGATCCCTTAATCAGGAATCAGCTCTCTCATCCCCCAGATGCCTATACCCCGAGGGA	1846
1678	QY	TCTATGAGATCTACCTCAGCTGCACAAGCCGGAAGACGTGAGGTGGCCCTAGTGGCT	1737
1847	DB	TCTATGAGATCTACCTCAGCTGCACAAGCCGGAAGACGTG-----	1887
1738	QY	GTGAGACCTGCTGAGTCCCATCGTTAGCTGTGGACCCCTGGCGTCTGTCTCACCCGC	1797
1888	DB	-----AGCTGTGGACCCCTGGCGTCTGTCTCACCCGC	1921
1798	QY	CAGTCACTCTGGCTATGGACCACTGTGGGAGCCCAAGCCCTGACAGCTGGAGCTGCGCC	1857
1922	DB	CAGTCACTCTGGCTATGGACCACTGTGGGAGCCCAAGCCCTGACAGCTGGAGCTGCGCC	1981
1858	QY	TCAAAAAGCAGTCTGTGCGAGGGCAGCTGGAGGACAGATGTGCTGCACCTGGCGAGGAG	1917
1982	DB	TCAAAAAGCAGTCTGTGCGAGGGCAGCTGGG---AGGATGTGCTGCACCTGGCGAGGAG	2038
1918	QY	CGCCCTCCCACTCTACTACTGTGCCAGCTGGAGGCCAGTGCCTGCTACGCTTCACCGAGC	1977
2039	DB	CGCCCTCCCACTCTACTACTGTGCCAGCTGGAGGCCAGTGCCTGCTACGCTTCACCGAGC	2098
1978	QY	AGCTGGGCGCTTTGCCCTGCTGGGAGAGCCCTCAGCTGTGGCTGCGGCCCAAGCGCTCA	2037
2099	DB	AGCTGGGCGCTTTGCCCTGCTGGGAGAGCCCTCAGCTGTGGCTGCGGCCCAAGCGCTCA	2158
2038	QY	AGCTGCTTCTGTTTGGCGCGGTGCCTGCACCTCCCTCGAGTACACATCCGGTCTACT	2097
2159	DB	AGCTGCTTCTGTTTGGCGCGGTGCCTGCACCTCCCTCGAGTACACATCCGGTCTACT	2218
2098	QY	GCCTGCATGACACCCAGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGG	2157
2219	DB	GCCTGCATGACACCCAGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGG	2278
2158	QY	GACAGCTGATCCAGGAGCCAGGCTCCTGCATTTCAAGGACAGTTACCAAACTGCGCC	2217
2279	DB	GACAGCTGATCCAGGAGCCAGGCTCCTGCATTTCAAGGACAGTTACCAAACTGCGCC	2338
2218	QY	TATCATCCAGATGTGCCAGCTCCCTGTGGAGAGATAGCTCCCTGTGAGTACCAGG	2277
2339	DB	TATCATCCAGATGTGCCAGCTCCCTGTGGAGAGATAGCTCCCTGTGAGTACCAGG	2398
2278	QY	AGATCCCTTTTATCAGATCTGGAAATGGCAGCAGCGGTACTTGCACTGCACCTTCACCC	2337
2399	DB	AGATCCCTTTTATCAGATCTGGAAATGGCAGCAGCGGTACTTGCACTGCACCTTCACCC	2458
2338	QY	TGGAGCGTGTACGCCCAAGCACTAGTGAACCTGGCGCTGCAAGCTGTGGGTGTGGCAGGTG	2397
2459	DB	TGGAGCGTGTACGCCCAAGCACTAGTGAACCTGGCGCTGCAAGCTGTGGGTGTGGCAGGTG	2518
2398	QY	AGGGCGACGGGAGAGCTTCAGCATCACTTCACATCACCAGGACACAGGTTTCTGT	2457
2519	DB	AGGGCGACGGGAGAGCTTCAGCATCACTTCACATCACCAGGACACAGGTTTCTGT	2578

9 JUL 1964

RESOL 6
AAV52940

ID AAV52940 standard; cDNA; 3014 BP.

XX
FBI

AAV52940; AC

XX

DT 25-MAR-2003 (revised)

DT 21-DEC-19

XX

DE Rat UNC-5

XX

KW UNC-5; UN

KW diagnosis

XX

OS
Rattus sp

XXVII

FH	Key	cdc
FF		

FEF

FYI

XX
DN
W09837089

XX
FN 00788605

27-AUG-1964

FD-302 (Rev. 4-15-64)

19-FEB-61

XX

PR 19-FEB-1968

XX

PA (REGC) U

XX

PI Tessier-1

XX

DR WPI; 1991

DR P-PSDB; 1

XX

AAV52941) and unc5h-2 (see AAV52943) cDNAs are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly from transfected host cells by utilising these vertebrate UNC-5 nucleic acids. The invention also provides unc-5 hybridisation probes and primers, vertebrate UNC-5-specific binding agents such as specific antibodies, and methods of making and using the subject compositions in diagnosis (e.g. genetic hybridisation screens for vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate vertebrate unc-5 gene expression), and in the biopharmaceutical industry (e.g. as immunogens, reagents for modulating cell guidance, reagents for screening chemical libraries for lead pharmacological agents, etc.). (Updated on 25-MAR-2003 to correct PI field.)

XX SQ Sequence 3014 BP; 596 A; 977 C; 849 G; 592 T; 0 U; 0 Other;

Query Match 77.6%; Score 2235.8; DB 2; Length 3014;
Best Local Similarity 88.5%; Pred No. 0;
Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;

QY 87 ATGGCCCTCGGCGCGGCTGTGGCCAGCGCTCTCGGCGATAGTCTCTCGCGCTTGGCTC 146
Db 1 ATGGCCCTCGGCGCGGCTGTGGCCAGTGTCTCTGGCGATAGTCTCTCGCGCTTGGCTT 60

QY 147 CGGCGCTCGGCTGCCAGCAGTGCACCGTGGCCAAACCGAGTGCCTGTGGCGCTTGGCT 206
Db 61 CGTGGTTGGGTCGCCAGCAGTGCACCGTGGCCAAACCGAGTGCCTGTGGCGCTTGGCT 120

QY 207 GACCTGCTTCCCTCCACTTCTCTGGTGGAGCGGAGATGTGTACATCTTCTTCAAGTGCACGGGAG 266
Db 121 GACCTGCTGCCCTACTCTCTGGTAGAGCTGTAGAGCGTGTACATTTGTCAAGAACAAAGCCG 180

QY 267 GTGCTGTTGTGTGCAAGGCGTGCCTGCCGCAACGAGATCTTCTTCAAGTGCACGGGAG 326
Db 181 GTGTGTGTGTGTGCAAGGCTGTGCTGCCACCCAGATCTTCTTCAAGTGCATGGGAA 240

QY 327 TGGGTGGCGGAGTGCACAGTGTGAGCGGAGCAGACAGCGGAGCAGTGTGGAGCCG 386
Db 241 TGGGTGGCGGAGTGCATACAGTAATTTGAACGAGCAGCAGCAGCAGCGGATTTGCCA 300

QY 387 ACCATGGAGTCCGCAATTAATGTCTCAAGCAGCAGTTCAGAGGTTCGGGCTGGAG 446
Db 301 ACCATGGAGTCCGTAATCAAGTATCAAGCAGCAGTTCAGAGGTTCGGGCTGGAG 360

QY 447 GAATATGTGTGCGAGTGTGCGTGGATGGAGCTCTCGGCGCAGCAGCAGCAGCAGAGCC 506
Db 361 GAATATGTGTGCGAGTGTGCGTGGATGGAGCTCTCGGCGCAGCAGCAGCAGCAGAGCC 420

QY 507 TACATCGGATACCGAGATTTGCGAAGATTTGAGCAGGAGCGCTGGCCAAAGGAGTG 566
Db 421 TACATCGGATTTGCCATTTTGGCAGAACTTTGAGCAGGAGCAGCAGCAGCAGGAAAGTG 480

QY 567 TCCCTGGAGCAGGGCATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 626
Db 481 TCACTGGAGCAGGGCATCTGTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540

QY 627 GTGGAGTGTCTCGGAACAGGAGCTGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 686
Db 541 GTGGAGTGTCTTCCGAATCAGGAGCTGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

QY 687 ACGCGGAGCAGCAGCTGTGTGTGCAAGGCGGCTTGTGTGACAGCGCCAACTACACC 746
Db 601 ACGCGGAGCAGCAGCTGTGTGTGCAAGGCGGCTTGTGTGACAGCGCCAACTACACC 660

QY 747 TGGTGGTGGCCAAAGCAATCTGTGCAAGTGTGCAAGGCGGCTGCTGCTGCTGCTGCTGCTGCT 806
Db 661 TGTGTGGCCAAAGCAATCTGTGCAAGTGTGCAAGGCGGCTGCTGCTGCTGCTGCTGCTGCT 720

QY 807 GTGAACGTTGGTGTGTGCAAGTGTGCAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 866
Db 721 GTGAACGTTGGTGTGTGCAAGTGTGCAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780

QY 867 GCTTGGCAGAAACGAGCGGAGTGTGCAACCCGCGGCTTCTCAACGGGCGGCTTCTTC 926
Db 867 GCTTGGCAGAAACGAGCGGAGTGTGCAACCCGCGGCTTCTCAACGGGCGGCTTCTTC 926

Db 781 GGCTGGCAGAAACGAGCGCGGAGCTGTGCAACCAACCCGGGCACTCTCTCAACGGGGCGCCCTTC 840
QY 927 TGTGAGGGGCGAAGATGTCCATGACCGCAGCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 986
Db 841 TGTGAGGGGCGAAGATGTCCAGAAACAGCCTGTGCGCACTCTG---TGCCCAAGTGGATGGG 897
QY 987 AGCTGGAGCCCGTGGAGCAAGTGTGGCCCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 1046
Db 998 AGCTGGAGTTCGTGGAGTAAGTGTGAGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 957
QY 1047 CGTGAAGTGTCTGAGCCAGCAGCCCGCAACGAGGAGGAGTGTGGAGGAGTGTGGAGGAGTGTGG 1106
Db 958 CGCAGTGTCTGTGACCCAGCAGCCCGCAATGAGGTGAGGAGTGTGGAGGAGTGTGGAGGAGTGTGG 1017
QY 1107 GACACCCGCAACTGTACAGTGAACCTCTGTGTGTACACAGTGTCTTCTGGCCCTCTGAGGAGCTG 1166
Db 1018 GACACCCGCAACTGTACAGTGAACCTCTGTGTGTACACAGTGTCTTCTGGCCCGAGGAGCTG 1077
QY 1167 GCCCTCTATGTGGCCCTCATGCGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 1226
Db 1078 GCTCTCTACATGGCCCTTGTGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 1137
QY 1227 ATCTCTGTTTATTTGCCGAAAGAGGAGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 1286
Db 1138 GGAATCATTTACTGTCCAGAGAGGAGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 1197
QY 1287 CTCACTTCAGGCTTCAGCCCGTGTAGCATCAAGCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1346
Db 1198 CTCACTTCAGGCTTCAGCCCGTGTAGCATCAAGCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1257
QY 1347 CTCACTTCAGGCTTCAGCCCGTGTAGCATCAAGCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1403
Db 1258 CTCACTTCAGGCTTCAGCCCGTGTAGCATCAAGCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1317
QY 1404 CGCAGGATGGGCGCCAGCCCAAGTTCCAGCTCACAATGGGCAAGCTGTGAGGCTGTGAGGCTGTG 1463
Db 1318 AGCAGGATGGAGCCAGCCCGTGTAGCATCAAGCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1377
QY 1464 GGTGGCGGCGCACAACATGTGCAACAGCTTCCCACTCTGAGGCGGAGGAGTGTGCTG 1523
Db 1378 GGGAGTGGCGCGCATACGTGTGCAACAGCTCACCACCTCTGAGGCTGTGAGGCTGTGCTG 1437
QY 1524 TCCCGCTCTCCACCCAGCAACTTCTCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1583
Db 1438 TCCCGCTCTCCACCCAGCAACTTCTCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1497
QY 1584 TATGGAGCTTCAACTTCTCGGGGCGGCTGTGATCCCTTAATACAGGATACAGGATACAGGCT 1643
Db 1498 TACGGAGCTTCAACTTCTCGGGGCGGCTGTGATCCCTTAATACAGGATACAGGATACAGGCT 1557
QY 1644 CTCACTCCCGCAGATGCCATATCCCGAGGAGAGATCTATGAGATCTATGAGATCTATGAGATCT 1703
Db 1558 CTCACTCCCGCAGATGCCATATCCCGAGGAGAGATCTATGAGATCTATGAGATCTATGAGATCT 1617
QY 1704 AAGCGGAGAGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1763
Db 1618 AAGCGGAGAGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1677
QY 1764 AGCTGTGAGACCCCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1823
Db 1678 AGCTGTGGGCGCCAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1737
QY 1824 GGGAGCCAGCCCTGTGAGCTGTGAGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 1883
Db 1738 GGGAGCCAGCCCTGTGAGCTGTGAGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 1797
QY 1884 TGGGAGCAGGATGTGCTGCACTGTGGCGAGGAGCGCCCTCCCACTCTACTACTGCTGCTGCTGCTG 1943
Db 1798 TGGG---AGGATGTGCTGCACTGTGGCGAGGAGTCACTTCCCACTCTACTACTGCTGCTGCTGCTG 1854
QY 1944 CTGAGGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2003
Db 1855 CTGAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1914

local cerebral glucose utilisation; phosphodiesterase 1-alpha; UNC5H1; calcium-independent alpha-latrotoxin receptor; CIRL; trke; synapsin 1A; epithelial discoidin domain receptor 1; synapsin 1B; neuroleptic; ss; tumour necrosis factor alpha; TNF-alpha; rat.

Rattus sp.

OS OS

Key Location/Qualifiers

CDS 1..2697

FT /tag= a

FT /product= "Rat netrin receptor UNC5H1"

PN WO200175440-A2.

XX PD 11-OCT-2001.

XX PF 02-APR-2001; 2001WO-GB001486.

XX PR 31-MAR-2000; 2000GB-00007880.

XX PR 26-MAY-2000; 2000GB-00012768.

XX PA (WELF-) WELFIDE CORP.

XX PI Cochran S, Paterson G, Ohashi Y, Morris B, Pratt J;

XX PI WPI; 2002-010813/01.

XX DR P-PSDB; AAU10543.

XX XX Novel chronic animal model of schizophrenia, useful for identifying anti-psychotic drugs and genes that are associated with schizophrenia.

XX PS Claim 1; Fig 8a; 79pp; English.

XX CC The invention relates to YSG polynucleotide fragments for use in diagnosing and/or developing treatments for schizophrenia using chronic animal models. The polynucleotides and their encoded polypeptides are used for identification of compounds which modulate the expression of YSG molecules, leading to the manufacture of schizophrenia medications. The sequences can also be used for testing candidate compounds for any effect on the polypeptides. Anti-schizophrenic effects of a compound can be determined by measuring local cerebral glucose utilisation (LCGU) or comparing its expression level with that of a control group. The sequences are useful in the identification of genes associated with schizophrenic states and in the development of an antibody. The sequences of the invention include phosphodiesterase 1-alpha, calcium-independent alpha-latrotoxin receptors (CIRL)-1,2&3 epithelial discoidin domain receptor 1 (trke), netrin receptor (UNC5H1), synapsins 1A and 1B and tumour necrosis factor (TNF) alpha. This sequence represents rat netrin receptor UNC5H1 (YSG7) DNA

XX SQ Sequence 2697 BP; 541 A; 864 C; 766 G; 526 T; 0 U; 0 Other;

Query Match 76.0%; Score 2189.4; DB 6; Length 2697;

Best Local Similarity 88.9%; Pred. No. 0;

Matches 2403; Conservative 0; Mismatches 291; Indels 9; Gaps 3;

Qy 87 ATGCGCGTCCGGCCCGGCGCTGTGGCCAGCGCTCTGGGCATAGTCTCTCGCGCTTGGCTC 146

Db 1 ATGGCGTCCCGCCCGGCGCTGTGGCCAGTGTCTCTGGGCATAGTCTCTCGCGCTTGGCTT 60

Qy 147 CGCGCTCGGCTGCCAGCAGAGTGCCACCGTGGCCAAACCCAGTGCCTGTGGTCCCAACCG 206

Db 61 CGTGGTTCGGGTGCCAGCAGAGTGCCACCGTGGCCAAATCCAGTGCCTGTGGTCCCAACCG 120

Qy 207 GACCTGCTTCCCGCTTCTCTGGTGGAGCCCGAGGATGTACATCGTCAAGAACAGCA 266

Db 121 GACCTGCTTCCCGCTTCTCTGGTGGAGCTTGGAGCGTGTACATTTGTCAAGAACAGCG 180

Qy 267 GTGCTGCTTGTGTGCAAGCGCGTCCCGCCAGCAGAGATCTTCTTCAAGTCAACGGGAG 326

Db 181 GTGTTGTTGGTGTGCAAGCGTGTGCTGCCACCCAGATCTTCTTCAAGTCAATGGGAA 240

Qy 327 TGGGTGCCCGAGGTGGACACGTGATCGAGCGCAGACAGCGGAGCAGTGGTGGCGG 386

2004 GAGGCCCTCAGCGTGTGGTGGCCCAAGCGCCTCAAGCTGCTTCTTGTGGCCGGTGGCC 2063

Db 1915 GAGGCCCTCAGCGTGTGGTGGCCCAAGCGCCTCAGGCTCTTCTGTGTGTCCCGTGGCC 1974

Qy 2064 TGCACCTCTCAGTACATCAACATCCGGTCTACTGCTGCATACACCCAGATGCATC 2123

Db 1975 TGTACGCTCCTTGAATACATCCAGTGTACTGTGCTTACACACCCAGCAGCTCTC 2034

Qy 2124 AAGGAGTGTGTGAGTGGAGAGAGCTGGGGGACAGCTGATGCAGGAGCCACCGGTC 2183

Db 2035 AAGGAGTGTGTGAGTGGAGAGAGCTAGTGTGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy 2184 CTGCACCTTCAAGGACATGATACCAACACCTGGGCTATTCATCCAGATGTCAGCTCC 2243

Db 2095 CTGCACCTTCAAGGACATGATACCAACACCTTACCTTCCATCCAGCAGCTGCCAGCTCC 2154

Qy 2244 CTGTGAGAGTAACTCTTGTGCTGCTACAGGAGATCCCTTTTATACATCTGGAAT 2303

Db 2155 CTGTGAGAGAGTAACTCTTGTGCTGCTACAGGAGATCCCTTTTACCATCTGGAAC 2214

Qy 2304 GGACCGCAGCGTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2363

Db 2215 GGACCGCAGCGTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2274

Qy 2364 GACCTGCGCTGCAAGCTGTGGGTGTGGAGTGTGGAGGCGAGCGGAGCTTTCAGCATC 2423

Db 2275 GACCTGCGCTGCAAGCTGTGGGTGTGGAGTGTGGAGGAGTGTGGAGGAGCTTTCACATC 2334

Qy 2424 AACTTCAACATCAACAGGACACAGGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2483

Db 2335 AACTTCAACATCAACAGGACACAGGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2394

Qy 2484 GGGTCTCCAGCGCTGTGGGCGCCAGTCCCTTCAAGATCCCTTCTCATTTCCGCGAAG 2543

Db 2395 GGGTCTCCAGCGCTGTGGGCGCCAGTCCCTTCAAGATCCCTTCTCATTTCCGCGAAG 2454

Qy 2544 ATAAATTCAGCTGAGCCACCTGTGTAGGCGGGTGGCGGAGCTGTGGCGGCTGTGGCGG 2603

Db 2455 ATCATCGCGCTGTGGAGCCACCTTGCAGCGCGGGCGCGAGTGTGAGAACTCTAGCCGAG 2514

Qy 2604 AACTCCAGCTGAGCGCATCTCAGCTTCTTGTCTTCCAGCCAGCCAGCCAGCCAGCCAG 2663

Db 2515 AACTCTCAGCTGAGCGCATCTTGTCTTCTTGTCTTCCAGCCAGCCAGCCAGCCAGCCAG 2574

Qy 2664 ATCTCTCAACCTGTGGAGCGCGGCTTCCCAACGCAACCTCAGCCAGCTGTGGTCA 2723

Db 2575 ATCTCTCAACCTGTGGAGCGCGGCTTCCCAACGCAACCTCAGCCAGCTGTGGTCA 2634

Qy 2724 GCAGTGGCTGAGTGGGCGGCGAGCAGCTGCTGCTTTCAGATGTGCGAGCTGAGTGC 2783

Db 2635 GCTGTGGCGGAGTGGGCGCAACAGATGCTGCTTTCAGCGGTGTGGAGGCGGAGTGT 2694

Qy 2784 TGAGCGCGCGCAGGCGCGG--ACACTTCACTCTTCAACAGCTTTGGCAGCCACCCAGGACA 2841

Db 2695 TGAGACCGAGCGGCGGATGCTTATGCTTATCTTCAACAGCTTTGACACTTGCAGGAGCA 2754

Qy 2842 GGCAGAACCGGACAGCGGCGGCTTCCCAACAGCGGCTTCCCAACAGCGGCTTCCCAACAGCGG 2878

Db 2755 GGCAGAACCGGACAGCGGCGGCTTCCCAACAGCGGCTTCCCAACAGCGGCTTCCCAACAGCGG 2791

RESULT 7

AA156843

ID AA156843 standard; cDNA; 2697 BP.

XX AC AA156843;

XX DT 14-FEB-2002 (first entry)

XX DE Rat netrin receptor UNC5H1 (YSG7) cDNA.

XX YSG; YSG7; schizophrenia; chronic animal model; LCGU; netrin receptor;

Db 241 TGGGTCCGCCAGTTCGATCACTAAATTGAACGACGACCGACAGCAGCGGATTCGA 300
QY 387 ACCATGAGGTCGCCAATTAAATGTTCTCAAGGACAGAGTCGAGAAGTGTTCGGCTGGAG 446
Db 301 ACCATGAGGTCGCCATCAAGTATCGAGGACAGAGGTAGAGAAGTGTTCGGCTGGAG 360
QY 447 GAATACCTGGTCCAGTGGTGGCATGAGCTTCCTCGGSCACCAACAAGAGTCAGAAGGCC 506
Db 361 GAATACCTGGTGGCATGAGTGGCATGAGCTTCCTCGGATACCAACAAGTCAAGAAGGCC 420
QY 507 TACATCCGATAGCAGATTGCGCAAGAACTTCAGCAGGAGCGCTGGCCCAAGAGGTG 566
Db 421 TACATCCGATAGTGGCATGAGTGGCATGAGCTTCCTCGGATACCAACAAGTCAAGAAGGTG 480
QY 567 TCCCTGAGCAGGACATGCTGCTGCCCTCGGTCACCGAGGAGCATCCCTCAGCCGAG 626
Db 481 TCACTGAGCAAGGATTTGATACCTTTGTGCCCCCCCCAGAAGGAATCCCCCAAGCTGAG 540
QY 627 GTGAGTGGCTCCGGAACGAGGACCTGTGGACCCGCTCCCTGGACCCCAATGTATACATC 686
Db 541 GTGAGTGGCTTCGAATGAGGACCTGTGGACCCCTCCCTCGATCCCAATGTGTATC 600
QY 687 ACCGGGAGCAGACCTGTGTGGTGGCAGACGCCGCTTGTGACACGGCCAACTACACC 746
Db 601 ACCGGGAGCAGACCTGTGTGGTGGCAGACGCCGCTTGTGACACGGCCAACTACACC 660
QY 747 TGGTGGCCAAAGAACATGTTGGACGCTGCGGACGCTGCGGAGCGCAGCTGTGGSCGC 866
Db 661 TGTGTGGCCNAGAACATGTTAGCCGCTGCGGAGGACCTCTCAGCGGTCATTTGTTAT 720
QY 807 GTGAACGGTGGTGGTGGACCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 866
Db 721 GTGAACGGTGGTGGTGGACCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 780
QY 867 GGTGTGCAGAAACGGAGCCGAGTGGACCAACCCGGCGCTCTCAACGGGGCGCTTTC 926
Db 781 GGTGTGCAGAAACGGAGCCGAGTGGACCAACCCGGCGCTCTCAACGGGGCGCTTTC 840
QY 927 TGTGAGGGGAGAAATGTCATGACCGACCGTCTCTCTCTCTCTCTCTCTCTCTCTCT 986
Db 841 TGTGAGGGGAGAAATGTCAGAAACAGCGCTGGCCCACTCTG---TGCCAGTGGATGGG 897
QY 987 AGCTGAGCCGCTGGAGCAAGTGGTGGGCTGGGCTGGACTGACCCCACTGGCGGAGC 1046
Db 898 AGCTGAGTTCGAGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 957
QY 1047 CGTGAAGTCTGACCCAGCAACCCGGAACGGAGGGAGGAGTGGCAGGSCATGACCTG 1106
Db 958 CGCGAGTGTCTGACCCAGCAACCCGCAATGGAGGTGAGGAGTGTGGGGTGTGACCTG 1017
QY 1107 GACACCCGAACTGTACAGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1166
Db 1018 GACACCCGAACTGTACAGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1077
QY 1167 GGCCTCTATGTGGGCTCTACGCGGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1226
Db 1078 GCTCTCTACATCGGCTGTGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1137
QY 1227 ATCTCTGTTTATGCGGAAGAGGAGGGGCTGGTCTCAGATGTGGTGGTGGTGGTGGT 1286
Db 1138 GGACTCATTTACTGTGCGAAGAGGAGGGCTGGACTCGGATGTGGCCGACTGTCTCATC 1197
QY 1287 CTCACCTCAGGCTTCAGCCGCTCAGCATCAGCCAGCAAGCAGCAACACCCCATCTG 1346
Db 1198 CTCACCTCAGGCTTCAGGCTGTGACATCAAGCCAGCAAGCAGCAACACCCCATCTG 1257
QY 1347 CTCACCATCAGCCGACCTCAGCACAC---ACCACCTACAGGGCAGTCTCTGTCC 1403
Db 1258 CTCACCATCAGCCAGCCTCAGCACACACCTACCATCAGGGGAGTCTATGTTG 1317
QY 1404 CGCAGGATGGGCCAGCCCAAGTTCCAGTCAACATGGGCACTGTCTCAGCCCTG 1463

Db 1318 AGCAGGATGGACCCAGCCCAAGTTCAGCTCTCTAATGGTCACTGTCTCAGCCCACTG 1377
QY 1464 GGTGGCGCGGCCACACACTGCAACCAAGCTCTCCACCTCTGAGGCCGAGAGTTCGTC 1523
Db 1378 GGGAGTGGCGGCCATACGTTGCAACCAAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437
QY 1524 TCCGCTCTCCACCCAGAACTACTTCGCTCCCTGCTGCCGAGGACCAAGCAATGACC 1583
Db 1438 TCCGCTCTCCACCCAAACTACTTTCTGCTGCCGCCGACCAAGCAATGACC 1497
QY 1584 TATGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGGCTC 1643
Db 1498 TAGGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACGGGATCAGCCTC 1557
QY 1644 CTATCCCCCAGATGCAATACCCCGAGGGAAGATCTATGAGATCTACCTCAGCTGCAC 1703
Db 1558 CTCATACCCCGGATGCCATCCCGAGGAAAGATCTACGAGATCTACCTCAGCTGCAC 1617
QY 1704 AAGCGGAAGACGTTGAGTTGCCCCCTAGCTGGCTGTGAGACCTCTGTGAGTCCCATGTT 1763
Db 1618 AAGCAGAAACGTTGAGTTGCCCCCTAGCTGGCTGTGAGACCTCTGTGAGTCCAGTCT 1677
QY 1764 AGCTGTGGACCCCTCGGCTCTGCTCACCGGCCAGTCACTCTGGCTATGAGACCTGTT 1823
Db 1678 AGCTGTGGGCCCCCAGGAGTCTCTCTCACCGGCCAGTCACTCTGCAATGGACCTGT 1737
QY 1824 GGGAGCCCAAGCCCTGACAGTGGAGCTGCGCTCAAAAAGCAGTGTGCGAGGCGAGC 1883
Db 1738 GGAGAGCCAGCCCTGACAGTGGAGTCTGCGCTCAAAAAGCAGTGTGCGAGGCGAGT 1797
QY 1884 TGGAGCAGAGTGTGCTGGACCTGGGCGGAGGAGGCGCTCCACCTCTACTACTGCGCAG 1943
Db 1798 TGGG---AGATGTGCTGACCTTGTGAGGAGTCACTTCCACCTCTACTACTGCGCAG 1854
QY 1944 CTGGAGCCAGTGTCTGCTACGCTTTCAGCGAGCTGGGCGGCTTGGCCCTGGTGGGA 2003
Db 1855 CTGGAGCCGGGCTCTGCTATGTTCAAGGAGCTGGGCGGCTTGGCCCTGGTAGGA 1914
QY 2004 GAGGCCCTCAGCTGCTGCGCCAAAGCGCTCAAAGTGTCTGTTTGGCGCGGTGGCC 2063
Db 1915 GAGGCCCTCAGCTGCTGCGCCAAAGCGCTCAAAGCGCTCAAAGCGCTCAAAGCGCT 1974
QY 2064 TGACCTCTCGCTCGATCAACATCCGGCTCTACTGCTGATGACACCAAGTGTGCTC 2123
Db 1975 TGTACCTCTGAGTACCAACATCCGAGTGTACTGCTTACACGACCCACGAGCTCTC 2034
QY 2124 AAGGAGTGGTGCAGCTGGAGAGCAGCTGGGGGAGAGTATCCAGAGCCACCGGTC 2183
Db 2035 AAGGAGTGGTGCAGCTGGAGAGCAGCTAGTGGAGAGCTGATCCAGAGCCCTCGGTC 2094
QY 2184 CTGCACTTCAAGACAGTTTACCAACCTGGGCTTATCCATCCAGATGTGCCAGCTCC 2243
Db 2095 CTGCACTTCAAGACAGTTTACCAACCTACCTACCTCTCCATCCAGAGCTCCGCTCC 2154
QY 2244 CTGTGAAGAGTAAAGCTCTCTGCTACGATCAGGAGATCCCTTTTATCAGTCTGGAAT 2303
Db 2155 CTGTGAAGAGCAAGCTACTTGTGAGTACAGGAGATCCCTTTTACCACATCTGGAAC 2214
QY 2304 GGCAAGCAGCGGTACTTGTGACCTGCACTTCACTTGGAGCGTGTGAGCCCGAGCTAGT 2363
Db 2215 GGCAAGCAGCGGTACTTGTGACCTTCACTTGGAGCGGATCAAGCCAGGACAGCAG 2274
QY 2364 GACCTGGCTTCAAGCTGTGGTGTGGCAGGTGGAGGGGAGCGGGCAGAGCTTCAGCATC 2423
Db 2275 GACCTGGCTTCAAGCTGTGGTGTGGCAGGTGGAGGGGAGATGGGAGAGCTTCAACATC 2334
QY 2424 AACTTCAACATCAACCAAGGACAGAGTGTGCTGAGCTGCTGGAGAGTGAAGCG 2483
Db 2335 AACTTCAACATCACTAAGGACAAAGGTTGTGTAATTTGTGGCTCTGGAGAGTGAAGG 2394
QY 2484 GGGGTCCAGCCCTGTGGTGGGCCCCAGTGTCAAGATCCCTTCTCTCTCTCTCTCTCT 2543
Db 2395 GGGGTCCAGCCCTGTGGTGGGCCCCAGTGTCAAGATCCCTTCTCTCTCTCTCTCTCTCT 2454

Db	247	GAGGTTGCCCTTACCTGGCTGTGAGACCCCTGCTGAGTGCCCATGTTAGCTGTGGACCCCC	306
Qy	1778	TGGCGTCTCTCTACACCCGGCAGTCATCCTGGCTATGGACCACTGTGGGAGAGCCACGCC	1837
Db	307	TGGGCTCTCTCTACACCCGGCAGTCATCCTGGCTATGGACCACTGTGGGAGAGCCACGCC	366
Qy	1838	TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGAGGACAGATGT	1897
Db	367	TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG--AGATGT	423
Qy	1998	GCTGCACCTGGCGAGAGAGGGCCCTCCACCTCTACTACTGCACGCTGAGAGGCCAGTGCC	1957
Db	424	GCTGCACCTGGCGAGAGAGGGCCCTCCACCTCTACTACTGCACGCTGAGAGGCCAGTGCC	483
Qy	1958	CTGCTACGTCCTTACCCAGCAGCTGGGCGCTTTGCCCTGGTGGGAGAGCCCTCAGCGT	2017
Db	484	CTGCTACGTCCTTACCCAGCAGCTGGGCGCTTTGCCCTGGTGGGAGAGCCCTCAGCGT	543
Qy	2018	GGCTGCCGCCAAGCGCTCAAGCTGCTTCTGTTTGGCCGGTGCCCTGCACTCCCTCGA	2077
Db	544	GGCTGCCGCCAAGCGCTCAAGCTGCTTCTGTTTGGCCGGTGCCCTGCACTCCCTCGA	603
Qy	2078	GTACAAATCCGGGCTTACTGCTTGATGACCCACCGATGCACTCAAGAGAGTGTGCA	2137
Db	604	GTACAAATCCGGGCTTACTGCTTGATGACCCACCGATGCACTCAAGAGAGTGTGCA	663
Qy	2138	GCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCAGGCTCCTGCACTTCAAGGA	2197
Db	664	GCTGGAGNAGCAGCTGGGGGACAGCTGATCCAGGAGCCAGGCTCCTGCACTTCAAGGA	723
Qy	2198	CAGTTACCAACCTGGCGCTATCCATCCACAGATGTGCCAGTCCCTGTGGAGAGTAA	2257
Db	724	CAGTTACCAACCTGGCGCTATCCATCCACAGATGTGCCAGTCCCTGTGGAGAGTAA	783
Qy	2258	GCTCCTTGTGAGCTACGAGGAGTCCCTTTTATCAGATCTGGAATGGCAGCAGCGGTA	2317
Db	784	GCTCCTTGTGAGCTACGAGGAGTCCCTTTTATCAGATCTGGAATGGCAGCAGCGGTA	843
Qy	2318	CTTGCACTGCACTTCACTTGGAGCGTGTGAGCCCGCAGCACTAGTGACCTGGCGTGC	2377
Db	844	CTTGCACTGCACTTCACTTGGAGCGTGTGAGCCCGCAGCACTAGTGACCTGGCGTGC	903
Qy	2378	GCTGTGGGTGTGGCAGTGTGAGGGGAGCGGCGAGAGCTTCAAGATCAACTCAACATCAC	2437
Db	904	GCTGTGGGTGTGGCAGTGTGAGGGGAGCGGCGAGAGCTTCAAGATCAACTCAACATCAC	963
Qy	2438	CAAGGACACAAGTTTGTGCTGAGCTGCTGGCTTGGAGAGTGAAGCGGGGTCCAGCGCT	2497
Db	964	CAAGGACACAAGTTTGTGAGCTGCTGGCTTGGAGAGTGAAGCGGGGTCCAGCGCT	1023
Qy	2498	GGTGGGCCCGAGTGCTTCAAGATCCCTTCTCATTTGGCAGAGAATAATTTCCAGCCT	2557
Db	1024	GGTGGGCCCGAGTGCTTCAAGATCCCTTCTCATTTGGCAGAGAATAATTTCCAGCCT	1083
Qy	2558	GGACCCACCTGTAGGGGGGGTGCCGATGTGGGAGACTCTGCCCGACAAATCTCCACCTGGA	2617
Db	1084	GGACCCACCTGTAGGGGGGGTGCCGACTTGGCGGACTCTGCCCGCAGAACTTCCACCTGGA	1143
Qy	2618	CAGCCATCTCAGCTTCTTTGCTTCCAAAGCCAGCCCGCAGCCGATGATCTCAACCTGTG	2677
Db	1144	CAGCCATCTCAGCTTCTTTGCTTCCAAAGCCAGCCCGCAGCCGATGATCTCAACCTGTG	1203
Qy	2678	GGAGGCGGGGCACTTTCCCGAACGGCAACTCAGCCAGCTGCTGCAGAGTGGGTGAGCT	2737
Db	1204	GGAGGCGGGGCACTTTCCCGAACGGCAACTCAGCCAGCTGCTGCAGAGTGGGTGAGCT	1263
Qy	2738	GGGCGACGACAGCTGGGCTCTTACAGTGTGAGAGCTGAGTGCTGAGGC	2789
Db	1264	GGGCGACGACAGCTGGGCTCTTACAGTGTGAGAGCTGAGTGCTGAGGC	1315

AAH99530
ID AAH99530 standard; cDNA; 1321 bp.
AC
AC AAH99530;
XX
XX 16-OCT-2001 (first entry)
XX
XX Human protein encoding cDNA sequence SEQ ID NO:365.
XX

Human; cancer; ulcer; HIV infection; human immunodeficiency virus; antiinflammatory; antirheumatic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac; central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antineoplastic; anaemia; antiaggregant; haemostatic; vulnery; antiulcer; osteopathic; eczema; dermatological; haemostatic; vulnery; antiulcer; osteopathic; eczema; dermatological; antiasthmatic; antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic; antiparkinsonian; infection; immunostimulant; gene therapy; antineuse therapy; vaccine; inflammation; antianginal; rheumatoid arthritis; septic shock; pancreatitis; cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity; genetic disease; haematopoietic disorder; platelet disorder; asthma; thrombocytopaenia; osteoporosis; severe combined immunodeficiency; allergic rhinitis; diabetes; multiple sclerosis; depression; Alzheimer's disease; Parkinson's disease; neurodegenerative disorder; neurological disorder; ss.

Homo sapiens.

WO200153455-A2.

26-JUL-2001.

22-DEC-2000: 2000WO-US035017.

23-DEC-1999: 99US-00471275.

21-JAN-2000; 2000US=00488723.
25-APR-2000; 2000US=00552317.

(HYCE-) HYCEO INC

U
3
T
YE
B

07/06/98 1000

07/06/98 1000

P-PSDB; AAM25589.

Isolated human pc

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

25963. The proteins can have activities based on the tissues and cells they are expressed in, such as: anti-inflammatory; antirheumatic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant; central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antinaemic; antidysregant; haemostatic; cardiant; antiulcer; osteopathic; dermatological; antiallergic; antisthmatic; antidabetic; cytostatic; neuroprotective; antidepressant; nootropic; antiparkinsonian; and immunostimulant. The proteins and polynucleotides encoding them can be used in gene therapy, antisense therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmunity, genetic diseases, haematopoietic disorders, anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers, osteoporosis, severe combined immunodeficiency, eczema, allergic rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzheimer's disease, Parkinson's disease, neurodegenerative and neurological disorders.

Sequence 1321 BP: 258 A; 440 C; 371 G; 252 T; 0 U; 0 Other;

Query Match		41.4%; Score 1193.6; DB 4; Length 1321;
Best Local Similarity		97.8%; Pred. No. 1.4e-210;
Matches 1295; Conservative		0; Mismatches 19; Indels 10; Gaps 8;
QY	1476	CACACACTGACACACAGCTCTCCACACTCTGAGGCGGAGGAGTTCGTCTCCCGCTCTCC 1535
DB	1	CACACACTGACACACAGCTCTCCACACTCTGAGGCGGAGGAGTTCGTCTCCCGCTCTCC 60
QY	1536	ACCAGAACTACTTCCGCTCCCTGCGGCGGAGGAGGACACATGATGATGAGGAGCTTC 1595
DB	61	ACCAGAACTACTTCCGCTCCCTGCGGCGGAGGAGGACACATGATGATGAGGAGCTTC 120
QY	1596	AATCTCTCGGGGCGGCTGATGATCCCTAATACAGATATGAGCTCTCTATCCCCCA 1655
DB	121	AATCTCTCGGGGCGGCTGATGATCCCTAATACAGATATGAGCTCTCTATCCCCCA 180
QY	1656	GATGCCATACCCGAGGAGAGATCTATGAGATCTACCTCAGCTTGCAAGCCGGAAGAC 1715
DB	181	GATGCCATACCCGAGGAGAGATCTATGAGATCTACCTCAGCTTGCAAGCCGGAAGAC 240
QY	1716	GTGAGTGTGCCCTAGCTGCTGTGACACCTGTGAGTCCATCGTTAGCTGTGACCC 1775
DB	241	GTGAGTGTGCCCTAGCTGCTGTGACACCTGTGAGTCCATCGTTAGCTGTGACCC 300
QY	1776	CCT-GGGCTCTGCTACCCGGGCGAGTATCCT-GGCTATGGACCACTGT-GGGGAGCCC 1832
DB	301	CCTGGGCTCTGCTTACCCGGGCGAGTATCCTGGGCTATGGACCACTGTGGGGAGCCC 360
QY	1833	AGCCCTGACAGCT-GGAGCCTGCGCTTCAAAAGCAGTCTGTGAGGCGAGCTGGGAGCA 1891
DB	361	AGCCCTGACAGCTGGGAGCTGCGCTTCAAAAGCAGTCTGTGAGGCGAGCTGGG--A 417
QY	1892	GGATGTGCTGCACTGGGCGGAGGAGGCGCTCCACCTCTACTACTGTCAGCTGGAGGC 1951
DB	418	GGATGTGCTGCACTGGGCGGAGGAGGCGCTCCACCTCTACTACTGTCAGCTGGAGGC 477
QY	1952	CAGTGTCTGCTGCTTCAACGAGCAGCTGGGCGGCTTTGCCCTGTGGGAGAGGCCCT 2011
DB	478	CAGTGTCTGCTGCTTCAACGAGCAGCTGAGCGGCTATGCCCTGTGGGAGAGGCCCT 537
QY	2012	CAGCGTGGCTGCGGCAAGCGCTCAAGCTGCTTCTGTTTGGCGCGGTGGCCCTGCACCTC 2071
DB	538	CAGCGTGGCTGCGGCAAGCGCTCAAGCTGCTTCTGTTTGGCGCGGTGGCCCTGCACCTC 597
QY	2072	CCTCGAGTCAACATCCGGGTCTACTGCTGCTGATGACACCCAGATGCACTCAAGAGGT 2131
DB	598	CCTCGAGTCAACATCTGCTACTGCTGCTGATGACACCTCAAGTGTGCACTCAAGTGT 657
QY	2132	GGTGAGCTGGAGAACAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCACTT 2191
DB	658	GGTGAGCTGGAGAACAGCTGCGGGACAGCTGATCCAGGAGCCACTGTACTGCACTT 717
QY	2192	CAAGGACGTATACCAACCTGGCTATCCATCCAGATGTGCCAGTCTCCCTGTGGAA 2251
DB	718	CAAGGACGTATACCAACCTGGCTATCCATCCAGATGTGCCAGTCTCCCTGTGGAA 777
QY	2252	GAGTAACTCTTGTGAGTACAGGAGATCCCTTTTATCATCTGAAATGGACGCA 2311
DB	778	GAGTAACTCTTGTGAGTACAGGAGATCCCTTTTATCATCTGAAATGGACGCA 837
QY	2312	GGGTACTTGCACTGACCTTCAACCTGGAGGCTGTGAGCTTGTGAGCTGCTGCTGC 2371
DB	838	GGGTACTTGCACTGACCTTCAACCTGGAGGCTGTGAGCTTGTGAGCTGCTGCTGC 897
QY	2372	CTGCAAGCTGTGGTGTGGCAGGTGAGGCGGAGGCGAGCTTCAAGATCAACTCAA 2431
DB	898	CTGCAAGCTGTGGTGTGGCAGGTGAGGCGGAGGCGAGCTTCAAGATCAACTCAA 957
QY	2432	CATCACCAGGACACAAGTGTGAGCTGCTGCTGAGAGTGAAGCGGGGTGCC 2491
DB	958	CATCACCAGGACACAAGTGTGAGCTGCTGCTGAGAGTGAAGCGGGGTGCC 1017
QY	2492	AGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCTATCGGAGAGATAATTTC 2551

DB	1018	AGCCCTGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAGATAATTTC 1077
QY	2552	CAGCTTGAGACCCACCTCTGTAGGCGGGGTGCCAGCTTGGCGGAGCTTGGCCAGAACTCCA 2611
DB	1078	CAGCTTGAGACCCACCTCTGTAGGCGGGGTGCCAGCTTGGCGGAGCTTGGCCAGAACTCCA 1137
QY	2612	CCTGAGCAGCATCTCAGCTTCTTTGCTCCAAAGCCAGCCCCACAGCCATGATCCTCAA 2671
DB	1138	CCTGAGCAGCATCTCAGCTTCTTTGCTCCAAAGCCAGCCCCACAGCCATGATCCTCAA 1197
QY	2672	CCTGTGGAGGCGGCGCACTTCCCAACGGCAACCTCAGCAGCTGGTGCAGCAGTGGC 2731
DB	1198	CCTGTGGAGGCGGCGCACTTCCCAACGGCAACCTCAGCAGCTGGTGCAGCAGTGGC 1257
QY	2732	T-GGACTGGGCCAGCCAGACGCTGGCCCTC-TTCACAGTG-TCCGAGGCTGAGTCTGAGG 2788
DB	1258	TGGGACTGGGCCAGCCAGACGCTGGCTTCTTTCACAGTG-TCCGAGGCTGAGTCTGAGG 1317
QY	2789	CCGG 2792
DB	1318	CCGG 1321
RESULT 11		
ID	AAS75738	standard; cDNA; 3646 BP.
XX	AAS75738;	
XX	AC	
XX	DT	13-FEB-2002 (first entry)
XX	DE	DNA encoding novel human diagnostic protein #11542.
XX	KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX	KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX	OS	Homo sapiens.
XX	PN	WO200175067-A2.
XX	PD	11-OCT-2001.
XX	PF	30-MAR-2001; 2001WO-US008631.
XX	PR	31-MAR-2000; 2000US-00540217.
XX	PR	23-AUG-2000; 2000US-00649167.
XX	PA	(HYSE-) HYSEQ INC.
XX	PI	Drmanac RT, Liu C, Tang YT;
XX	DR	WPI; 2001-639362/73.
XX	DR	P-PSDB; ABG11551.
XX	PT	New isolated polynucleotide and encoded polypeptides, useful in
XX	PT	diagnostics, forensics, gene mapping, identification of mutations
XX	PT	responsible for genetic disorders or other traits and to assess
XX	PT	biodiversity.
XX	PS	Claim 1; SEQ ID NO 11542; 103pp; English.
XX	CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX	CC	sequences. (I) is useful as hybridisation probes, polymerase chain
XX	CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX	CC	and in recombinant production of (II). The polynucleotides are also used
XX	CC	in diagnostics as expressed sequence tags for identifying expressed
XX	CC	genes. (I) is useful in gene therapy techniques to restore normal
XX	CC	activity of (II) or to treat disease states involving (II). (II) is
XX	CC	polypeptide in tissue, as molecular weight markers and as a food
XX	CC	supplement. (II) and its binding partners are useful in medical imaging
XX	CC	of sites expressing (II). (I) and (II) are useful for treating disorders

QY	544	AGGAGCCGTGGCCAAAGGAGGTGTCCTCTGGACGAGGCATCGTGTGTCGCCCTGCCGTCAC	603
Db	528	AGGAGCCTCTGGGCAAGGAGGTGCCCCCTGGACCATGAGTTCTCTGCAGTGCCTGCCGCGC	587
QY	604	CGGAGGCATCCCTCTCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGTTGACCCGT	663
Db	588	CGAGGGGGTGCCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCTATCGACCCCA	647
QY	664	CCCTGGACCCCAATATATCATACGCGGGAGACAGCTGGTGGTGGACAGGCGCCGCC	723
Db	648	CCCAGGACCAAACTTCTCTGCTCACCATCGACCAAACTCATCATCGCCAGGCGCCGCC	707
QY	724	TTGCTGACAGGCCCAACTACCTCGCTGGCCAAAGAACATGTTGGACAGTGCCTGCCGACGG	783
Db	708	TGTGGGACACTGGCAACTATACCTGCTGGCCAAAGAACATCTGGGCAAAAGCCGCGAGCA	767
QY	784	CCCTCCGTGCTGTATCGTCTACGTGAAAGCGTGGGTGGTTCGACGTGGACCCAGATGGTCCG	843
Db	768	CCACTGCCACGGTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCGAGATGGTCCAC	827
QY	844	CTTGACGCGCAGCTGTGGGCGCGGTGGCAGAAACGAGCGCGAGCTGCACCAACCCGG	903
Db	828	CCCTGCTCAACCGCTGTGGCGAGGTGGCAGAAGCGCACCCGAGCTGCACCAACCCGG	887
QY	904	CGCCTCTCAACGGGGCGCTTCTGTGAGGGGAGAAATGTCATGACCCGACCGCTCTCCT	963
Db	888	CTCCACTCAACGGAGGGCGCTTCTGCGAGGGCCAGGCAATCCAGAA--GACCGCTGCA	944
QY	964	CTCTGCTTGTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTGGGCTGTGGGC	1023
Db	945	CCACCATCTGCCAGTCATGGGCGTGGACGGAGTGGAGCAAGTGGTCAGGCTCGACCA	1004
QY	1024	TGGACTGACCCACTGGCGGAGCCGTGAGTCTCTGACCCAGCACCCCGCAACGGAGGG	1083
Db	1005	CTGAGTGTGCCCACTGGCGTAGCCGAGTGCATATGGCGCCGCCACCCCAAGACGGAGGC	1064
QY	1084	AGGAGTGCACGGGCACTGACCTGGACACCGCAACTGTACAGTGAACCTGTGTGACACA	1143
Db	1065	GTGACTGACGGAGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAAC	1124
QY	1144	GTGCTCTTGGCCCTGAGACGTGGCCCTCTATATGTTGGGCTCTATCGCGGTGGCGGTGCGC	1203
Db	1125	TGGAGGCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGTGGGCACTCTTGTGGTGG	1184
QY	1204	TGGTCTGCTGTGCTGTCTCTATCTCGTTTATTTGCGGAAGAGGAGGGGCTGGACT	1263
Db	1185	TGGCAATCCTCATGGCGTGGGGGTGGTGTGATACGCCGCAACTGCCGTGACTTCGACA	1244
QY	1264	CAGATGTGGCTGACTCGTCCATT---CTCACCTCAGGCTTCCAGCCGCTCAGCATCAAGC	1320
Db	1245	CAGACATCACTGACTCATCTGCTGCCCTGACTGGTGGTTCACCCCGTCAACTTTAAGA	1304
QY	1321	CCAGCAAGACGACAAACCCCATTTGCT-----CACCATCTCAGCGGAGCTCAGACCA	1374
Db	1305	CGGCAAGGCCCAAGTAACCGCAGTCTCTACACCCCTCTGTGGCTCTGACTCTCAGACCA	1364
QY	1375	CCACCACTTACAGGGGAGTCTCTGTCCCGGAGGATGGCCCGAGCCCAAGTTCACAGC	1434
Db	1365	GGCGCGGATCTACCGCGGACCCGTTATGCCCCGTGAGGACTCCACGACAAATCCCCA	1424
QY	1435	TCACCAATGGGCACTGTCTAGCCCC-----TCACCTCAGGCTTAAAGGCTCAAGGCTCTACAGTCCCA	1484
Db	1425	TGACCACTCTCTCTGTCTGGACCCCTTACCGAGCCTTAAAGGCTCAAGGCTCTACAGTCCCA	1476
QY	1461	-----CTGGGTGGGGCGCGC	1544
Db	1485	GCACCAAGGGCTCTGGGCGCAGGCTTGGCAGATGGGGCTGACCTGCTGGGGGTCTTGGCGC	1544
QY	1477	ACACATGACACACAGCTCTCCACCTCTGAGGGCCGAGGAGTGTGCTCCCGGCTCTCCA	1536
Db	1545	CTGGCACAATCCCTAGCGATTTTGGCCCGGACACCCACTTCTCTGCACCTGCGCAGCGCCA	1604
QY	1537	CCAGAACTAC-----TTCCCGCTCCCTGCCCCGAGGACACAGCAATGACTTATGGGA	1590

Db	1605	GCCTCGGTTCCACGACAGTCTTTGGGCTCGCCCGAGACCCAGGGAGACGCTCAGCGCA	1664
Qy	1591	CGTTCAACTTCCTCGGGGCCCGCTGATGATCCCTTAATACAGGTATCAGCCTCCTCATCC	1650
Db	1665	CGTTGGCTCGCTGGGTGGAGGCTCAGCATCCCCGGCACAGGGGTGAGTTGCTGGTGC	1724
Qy	1651	CCCCAGATGCCATACCCCGAGGAGAGATCTTAGATCTACCTCAGCTCAGCTGCACAGCCGG	1710
Db	1725	CCAATGGAGCAATCCCAAGGCAAGTTCTACGAGATGTATCTACTCATCAACAAGGGCAG	1784
Qy	1711	AAGACGTGAGTTGGCCCTTAGCTGGCTGTCAGACCTGCTGAGTCCCATCGTTAGCTGTG	1770
Db	1785	AAAGTACCTTCGCGCTTTCAGAAGGACCCAGACAGTATTGAGCCCTCGGTGACCTGTG	1844
Qy	1771	GACCCCTGGCGTCTGCTCA CCGCGCCAGTCACTCTGGCTATGAGCACTGTGGGGAGC	1830
Db	1845	GACCA CAGGCTCTCTGCTGTGCGGCCGTGATCTCTACCATGCGCCCATGTGCGGAG	1904
Qy	1831	CGACCTCAGACGCTGGAGCTCGGCTCAAAAGCAGTCGTGCGAGGCACTGGGAGC	1899
Db	1905	TCAGTGCCCGTGA CTGATCTTTACGCTCAGACCAGGCCCCACAGGGCCACTGGGAG-	1963
Qy	1891	AGGATGTCTGCACCTGGGCGAGGAGCGCCTCCACCTCTACTCTACCTGCGAGCTGGAGG	1950
Db	1964	--GAGTGGTGA CCGCTGGATGAGAGACCTGAACACACCTGCTACTGCCAGCTGGAGC	2021
Qy	1951	CCAGTGCCTGCTAGCTTCA CCGAGAGCTGGGCGCGCTTTGCCCTGGTGGGAGAGGCC	2010
Db	2022	CCAGGCTCTGCATCTCTCTGACAGCTGGGCACTACGTTTCA CCGGCGAGTCTT	2081
Qy	2011	TCAGGTGGTTCGCGCAAGCGCTCAAGCTGCTTCTGTTTGGCGCGGTGGCTGCACCT	2070
Db	2082	ATTCGCGCTCAGCAGTCAAGGGCTCCAGTGGCCGCTTCGCGCCGCCCTCTGCACT	2141
Qy	2071	CCCTCGAGTACAACATCCGGTCTACTGCTGATGACACCCAGATGCACTCAAGGAGG	2130
Db	2142	CCCTGGAGTACAGCTCCGGTCTACTGCTCGTGAGGACAGCGCTGTAGCACTGAAGGAGG	2201
Qy	2131	TGGTGCAGCTGGAGAACGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCACT	2190
Db	2202	TGCTGGAGCTGGAGCGGACTCTGGCCGATCTTGGTGAGAGAGCCGAACCGCTAATGT	2261
Qy	2191	TCAAGGACAGTTACACAACCTGGCCCTATCCATCCAAGATGTCGCAGCTCCCTGTGGA	2250
Db	2262	TCAAGGACAGTTACAACAACCTGGCCCTTCCCTCCAATGACCTCCGCCCATTTGGA	2321
Qy	2251	AGAGTAAGCTCCTGTGACTACAGAGATCCCTTTTATCATCTGGAATGGACAGC	2310
Db	2322	GGAGCAAGCTCTGCGCAATACAGAGAGATCCCTTCTATCATATTGGAGTGGAGCC	2381
Qy	2311	AGCGTACTTCACTGCACTTCACTCGGAGCGTGTGAGCCCGAGCACTAGTGCCTGG	2370
Db	2382	AGAAGGCCCTCACTGCACTTTCACCTGGAGAGCACAGCTTGGCCCTCCACAGAGCTCA	2441
Qy	2371	CCTGCAAGCTGTGGTGTGGAGGTGGAGGGCGAGGGCAGAGCTTCAGCATCAACTTCA	2430
Db	2442	CCTGCAAGATCTGCTGCGCAAGTGGAAAGGGGAGGGCCAGATATTCAGCTGCATACCA	2501
Qy	2431	ACATCACCAG--GACACAAGTTTGTGAGTCTGGCTCTGGAGAGTGAAGCGGGG	2487
Db	2502	CTCTGGCAGAGACACTGCTGGCTCCCTGGACACTCTCTGCTCTGCGCCCTGGCAGCACTG	2561
Qy	2488	TCCCAGCCCTGGTGGGCCCGCTGCTTCAAGATCCCTTCTCATTTGGCAGAGAGATAA	2547
Db	2562	TCACCACCGAGCTGGGACCTTATGCTTCAAGATCCCACTGCTCATCCGCCAGAGATAT	2621
Qy	2548	TTTTCAGCTGCAACCACTCTAGGCGGGTGGCGACTGGCGGACTCTGGCCCCAGAAAC	2607
Db	2622	GCAACAGCCTAGATGCCCCCCAACTCACGGGGCAATGACTGGGGATGTTAGCACAGAGC	2681
Qy	2608	TCCACTCTGACAGCCATCTCAGCTTCTTTGGCTCCAGGCCAGGCCCAAGCAATGATCC	2667

[illegible]

1185 TGGCAATCTCATGCGGCTGGGGTGGTGTACCGCCGCACTGCGGTGACTTCGACA 1244
 1264 CAGATGTGGTGAATCGTCCATT---CTACCTCAGGCTTCCAGCCCGTCAAGCATCAAGC 1320
 1245 CAGACATCACTGACTCATCTGCTGCCCTGACTGGTGTTCACACCCGCTCAACTTAAAG 1304
 1321 CCAGCAAGCAGACAAACCCGCTCTGTGT-----CACCATCCAGCGGACCTCAGCACCA 1374
 1305 CGGCAAGGCCAGTAACCCGCGAGCTCTACACCCCTCTGTGCTCTGACCTGACACCA 1364
 1375 CCACCACTTACAGGCGAGTCTCTGCTCCCGGAGGATGGCCCGACCCCAAGTTCAGC 1434
 1365 GCGCCGCGCATCTACCGCGGACCCGCTGTATGCCCTGAGGACTCAACCGACAAATCCCCA 1424
 1435 TCACCAATGGGCACTGCTCAGCCCC-----1460
 1425 TGACCACTCTCTCTGCTGAGACCCCTTACCAGCCTTAAAGTCAAGGTCTACAGTCCA 1484
 1461 -----CTGGGTGGCGCGCC 1476
 1485 GCACCAAGGCTCTGGGCCAGGCTTGCGAGATGGGGCTGACCTGCTGGGGGTCTTGCGC 1544
 1477 ACACACTGCACACAGCTCTCCACCTCTGAGGCGGAGGAGTTCGTCTCCCGCTCTTCCA 1536
 1545 CTGGCAATACCTTAGGATTTGCGCCGGGACACCCACTCTCTGCACTGGGAGCGCCA 1604
 1537 CCCAGAACTAC-----TTCGGTCTCCCTGCGCCGAGGACCCAGCAACATCACTATCGGA 1590
 1605 GCCTCGGTTCCAGCAGCTCTTGGGCTCTGCGCCGAGACCCAGGAGCAGGCTCAGCGCA 1664
 1591 CTTTCAACTCTCTCGGGGCGGCTGATGATTCCTTAATACAGGTATCAGCTCTCTATCC 1650
 1665 CTTTGGCTGCTGGTGGGAGGCTCAGCATCCCGCCGACAGGGGTGAGTCTGCTGTGC 1724
 1651 CCCAGATGCATACCCCGAGGGAAGATCTATGATCTACTCTCAGCTGCACAAAGCGG 1710
 1725 CCAATGAGCCATTCCTCCAGGCGAAGTCTACGAGATGTATCTACTCATCAACAGGAG 1784
 1711 AAGAGTGGAGTTGCCCTAGCTGGCTGTGACACCTCTGAGTCCCATGCTTACCTGTG 1770
 1785 AAGTACCTGCGCTTTCAAGAGGAGCCAGACAGTATTGAGCCCTTCGGTGAACCTGTG 1844
 1771 GACCCCTGGGCTCTGCTCACCCTGCGGAGTCTATCTGCTGATGACACCTGTCGGGAGC 1830
 1845 GACCCACAGGCTCTGCTGTGCGCCGCGCTCATCTCACCATGCCCCACTGTGCCGAAG 1904
 1831 CCAGCCTTGACAGTGGAGCTGCGCTCAAAAAGCAGTCTGTGGAGGCGAGCTGGGAGC 1890
 1905 TCAGTCCCGTGAATGATCTTTTACCTCAAGACCCAGGCCACCCAGGCGCTGGAG- 1963
 1891 AGGATGTCTGACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCGAGCTGGAGG 1950
 1964 --GAGTGTGACCTGGATGAGGAGACCTTGAAACACACCTCTGCTACTGCGAGCTGGAGC 2021
 1951 CCAGTGCCTGTACGTCTTACCGAGCAGCTGGGCGGCTTTGCCCTGTGGGAGGAGGCC 2010
 2022 CCAGGCGCTGTCACTCTCTGTGACCACTGGGCGACCTAGCTGTTCAGGCGGAGTCTCT 2081
 2011 TCAGCTGGCTGCGGCCAGGCGCTCAAGCTGCTTCTGTTTGGCGCGGTGGCTGCACCT 2070
 2082 ATTCCGCTGAGCAGTCAAGCGGCTCAGCTGGCGCTTTCGCCCGCCCTCTGCACCT 2141
 2071 CCCTGAGTCAACATCCGGGTCTACTGCTGCTGATGACCCCAAGTGCATCAAGAGG 2130
 2142 CCCTGAGTCAAGCTCCGGGTCTACTGCTGGAGGACACGCTCTGAGCTGAAGAGG 2201
 2131 TGGTGCAGTCAAGAGCAGCTGGGGGAGCAGCTGATCCAGGAGCCAGGGTCTGCACT 2190
 2202 TCGTGGAGTGGAGCGGATCTTGGGCGGATATCTTGGTGGAGGAGCGGAAACCGCTAATGT 2261
 2191 TCAAGGACAGTTTACACACACTGCGCTATCCATCCAGATGTGCCAGCTCCCTGTGGA 2250

2262 TCAAGGACAGTTTACCACAACTTGGCGCTCTCCCTCCATGACCTCCCCATGCCCATTTGA 2321
 2251 AGAGTAAGCTCTGTGTCAGTACCAGGAGATCCCTTTTATCACATCTGGAATGGCAGCG 2310
 2322 GGAGCAAGCTGTGGCCAAATACAGGAGATCCCTTCTATCACATTTGGAGTGGCAGCC 2381
 2311 AGCGGTACTTGCACCTGACCTTACCTTGGAGGCTGTAGCCCCAGCACTAGTGAAGCTGG 2370
 2382 AGAAGGCCCTCCACTGCTTTCACCTTGGAGGCGACAGCTTGGCTCCACAGAGTCA 2441
 2371 CTTGCAAGCTGTGGGTGTGGAGTGGAGGCGACGGGAGAGCTTCAGCATCACTTCA 2430
 2442 CTTGCAAGATCTGCTGCGGCAAGTGAAGGAGGAGGCGGAGATATTCAGCTGCATACCA 2501
 2431 ACATCAACCAAG---GACACAAGCTTTTGGAGCTGTGGCTCTGGAGAGTGAAGCGGGG 2487
 2502 CTTGGCAGAGACACCTGCTGGCTCCCTGGAGACTCTCTGCTCTGCCCCCTGGCAGCACTG 2561
 2488 TCCAGCCCTGTGGGCCCCAGTGCCTTCAAGATCCCTTTCCTCATTTGGCAGAGATAA 2547
 2562 TCACCAACCCAGCTGGGACCTTATGCTTCAAGATCCCACTGTCCATCCGCCAGAGATAT 2621
 2548 TTTCCAGCTGGAGCCACCTGTAGGGGGGTGCGGAGTCCGAGTCTGCGCCAGAAAC 2607
 2622 GCAACAGCTTAGATGCCCCCACTCAGCGGGCAATGACTGGCGGATGTTAGCAGAGAGC 2681
 2608 TCCACCTGGAGACGACCTCTCAGCTTCTTTCCTTCCAAAGCCAGCCCCACAGCATGATCC 2667
 2682 TCTCTATGGACCGGTACTGAAATATTCTTCCACCAAGGAGGCCACCGGTTGATCC 2741
 2668 TCAACCTGTGGGAGGCGGCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCAGAG 2727
 2742 TGGACCTCTGGAAGCTCTGCAGCAGGACGATGGGACCTCAACAGCTGGCGAGTGCCT 2801
 2728 TGGCTGAGCTGGGCGCAGCAGCGTGGCTCTTTCAGTGTGCGAGGCTGAGTGTGA 2786
 2802 TGGAGGAGATGGGCAAGAGTGAATGCTGTGGCTGTGGCCACCGAGCGGGACTGCTGA 2860

RESULT 14

ABK92062
 ID ABK92062 standard; DNA; 2995 BP.

AC ABK92062;

XX
 XX
 DT 14-AUG-2002 (first entry)

XX
 XX
 DE DNA encoding novel UNC5 receptor-like protein.

XX
 DE Cancer; metabolic disorder; skin disorder; infectious disease; anorexia;
 behavioral disorder; valve disease; endocrine disorder; heart disorder;
 blood disorder; anxiety disorder; brain disorder; inflammatory disorder;
 neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 immune disorder; haematopoietic disorder; dyslipidaemia;
 metabolic disturbance; obesity; metabolic syndrome X; wasting disorder;
 gene therapy; transgenic animal; human; gene; ds.

XX Homo sapiens.

XX WO200229058-A2.

XX 11-APR-2002.

XX 05-OCT-2001; 2001WO-US031248.

XX 05-OCT-2000; 2000US-0238323P.

PR 05-OCT-2000; 2000US-0238325P.

PR 06-OCT-2000; 2000US-0238372P.

PR 06-OCT-2000; 2000US-0238373P.

PR 06-OCT-2000; 2000US-0238379P.

PR 06-OCT-2000; 2000US-0238382P.

PR 06-OCT-2000; 2000US-0238383P.

PR 06-OCT-2000; 2000US-0238384P.

Db 1401 AGCCAGCGCGGATCTACCGGACCCGCTGTATGCTGAGGACTCCACCGACAAT 1460
QY 1430 CCAGCTCACCATGGGACCTGCTAGCCCC----- 1460
Db 1461 CCCATGACCACTCTCTCTGCTGAGACCCCTTACCAGCCTTAAAGTCAAGGTCTACAG 1520
QY 1461 -----CTGGTGGCG 1471
Db 1521 CTCAGCACACGGGCTCTGGSCCAGGCTGGCAGATGGGCTGACCTGCTGGGCTTT 1580
QY 1472 CGCCACACACTGACACAGCTCTCCACCTCTGAGGCGGAGGAGTTCGTCTCCGCGCT 1531
Db 1581 GCGCCTGGCAGATACCTAGCAATTCGCGGAGACCCACTTCTGACCTGGGCGAG 1640
QY 1532 CTCACCCC-----AGAACTACTTCCGCTCCCTGCCCCGAGGACACAGCAATGACCTA 1585
Db 1641 CGCCAGCTCGGTCCAGCAGCTCTTGGSCCTCGCCGAGACCCAGGAGCAGCGCTAG 1700
QY 1586 TGGGACCTTCAACTTCTGCGGGGCGGCTGATGATCCCTTAATACAGGTATCAGCCTCT 1645
Db 1701 CGGCACTTTGGCTGCTGGGTGGGCTCAGCATCCCCGCGCACAGGTGTGAGCTTGCT 1760
QY 1646 CATCCCCCAGATGCCATACCCGAGGAAAGATCTATGAGATCTACCTCAGCTGCACAA 1705
Db 1761 GGTGCCCAATGAGAGCCANTCCAGGGCAAGTCTACGAGATGTATCTACTCATCACAA 1820
QY 1706 GCGGAGAGAGCTGAGGTGGCTTGGCTTAGCTGCTGACACCTGCTGAGTCCCATCGTTAG 1765
Db 1821 GGCAGAAAGTACCTCGCGCTTTTCAAGAGGACCCAGACAGTATTGAGCCCTCGGTGAC 1880
QY 1766 CTGTGACCCCTCGGCTGCTGCTCACCAGGAGTATCTGCTGCTATGACCACTGTGG 1825
Db 1881 CTGTGACCCACAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1940
QY 1826 GAGCCAGCCCTGACAGCTGGAGCTGGCGCTCAAAAAGCAGCTGCGGAGGAGCAGCTG 1885
Db 1941 CGAAGTCAGTCCGCTGACGTGATCTTTTACGTCAAGCCAGCCAGCCAGGAGCCACTG 2000
QY 1886 GGAGCAGGATGTGTCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCT 1945
Db 2001 GGAGCAGGAGTGTGACCTGGATGAGGAGACCTGAAACACACCTCTACTGCCAGCT 2060
QY 1946 GGAGCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2005
Db 2061 GGAGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2120
QY 2006 GGCCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2065
Db 2121 GTCTATTCCGCTCAGCAGTCAAGCGCTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2180
QY 2066 CACCTCCTCGAGTACAAATCCGCGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2125
Db 2181 CACCTCCTCGAGTACAGCTCCGCGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2240
QY 2126 GGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2185
Db 2241 GGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2300
QY 2186 GCATTTCAAGACAGTTTACCAACCTCGCTATCCATCCACGATGTGCCAGCTCCCT 2245
Db 2301 AATGTTCAAGACAGTTTACCAACCTCGCTATCCATCCATCCATCCATCCATCCATCCATCCAT 2360
QY 2246 GTGGAAGATGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2305
Db 2361 TTGAGAGGAGAGCTGCTGGGCAATACAGAGATCCCTTCTATCATCATTTGGAGTGG 2420
QY 2306 CAGCAGCGGTACTTGCATCGACCTTCCCTGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2365
Db 2421 CAGCCAGAGAGGCTTCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2480
QY 2366 CCGTGGCTCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2425
Db 2481 GCTCACCTGCAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2540

QY 2426 CTTCAACATCACCAAGGACA---CAAGGTTTGTGAGCTGCTGCTCTCGAGAGTGAAGC 2482
Db 2541 TACCACCTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCTGGCAG 2600
QY 2483 GGGGGTCCAGAGCCCTGGTGGGCCCCAGTCCCTTCAAGATCCCTTCTCTCATTTGGGAGAA 2542
Db 2601 CACTGTCAACCCAGCTGGGACCTTATGCTTCAAGATCCCACTGTCCATCCGCCAGAA 2660
QY 2543 GATAATTTCCAGCTGACCCACCTGTAGGGGGTGGCGACTGGCGGACTCTGGCCCA 2602
Db 2661 GATATCAACAGCCTAGATGCCCCAACTCAAGGGGAATGACTGGCGGATGTAGCACA 2720
QY 2603 GAACTCCACCTTGCAGAGCATCTCAGCTTCTTTTGGCTTCCAGGCCAGGCCCCAGCCAT 2662
Db 2721 GAAGCTCTCTATGAGACGGTACCTGAATTAATTTGCCACCAAGCGAGCCCAAGGGTGT 2780
QY 2663 GATCCTCAACCTTGGGAGCGCGGACCTTCCCAAGCGCAACCTTCAAGCAGCTGGCTGC 2722
Db 2781 GATCCTGGACCTCTGGGAGCTCTGAGCAGGACGATGGGAGCTCAACAGCCTTGGCGAG 2840
QY 2723 AGCAGTGGCTGAGTGGGCGAGCCAGCAGCTGGCCTTTTCCAGAGTGTGCGAGGCTGAGTG 2782
Db 2841 TGCCTTGGAGAGATGGCAAGAGTGAATGCTGGTGGCTGTGCGCCACCGACGGGACTG 2900
QY 2783 CTGAGGC 2789
Db 2901 CTGAGCC 2907

RESULT 15
ABQ93898
ID ABQ93898 standard; DNA; 2895 BP.
XX
AC ABQ93898;
XX
DT 01-NOV-2002 (first entry)
XX
DE Human transmembrane receptor UNC5H2-like NOV11 DNA, SEQ ID NO:37.
XX
KW Human; NOVX; neurological disorder; Alzheimer's disease;
KW Huntington's disease; Parkinson's disease; pain; behavioural disorder;
KW addition; tuberculous sclerosis; cancer; immune disorder; allergy;
KW autoimmune disease; myasthenia gravis; asthma; arthritis; diabetes;
KW thyroiditis; cardiovascular disease; hypertension; reproductive disorder;
KW endometriosis; incontinence; psoriasis; scleroderma; alopecia; ulcer;
KW pancreatitis; cirrhosis; glomerular endotheliosis; bacterial infection;
KW polycystic kidney disease; endocrine disorder; obesity; cardiomyopathy;
KW atherosclerosis; cell signal processing-related disorder;
KW metabolic pathway regulation disorder; cytostatic; neuroprotective;
KW antiinflammatory; immunosuppressive; analgesic; antithrombotic;
KW dermatological; antibacterial; antiarthritic; hepatotropic; neurogenesis;
KW differentiation; proliferation; motility; haematopoiesis; wound healing;
KW angiogenesis; forensic biology; transgenic animal; drug screening;
KW gene therapy; NOV11; transmembrane receptor UNC5H2-like; chromosome 10;
KW gene; ds.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
5'UTR 1..30
FT /*tag= a
FT 31..2868
FT /*tag= b
FT /*tag= "NOV11"
FT /product= 31..108
FT /*tag= c
FT 109..2865
FT /*tag= d
FT /product= "Mature NOV11 protein"
FT 2869..2895
FT /*tag= e
XX

WO200253742-A2.
 11-JUL-2002.
 07-JAN-2002; 2002WO-US000375.
 05-JAN-2001; 2001US-0260018P.
 08-JAN-2001; 2001US-0260360P.
 28-FEB-2001; 2001US-0272411P.
 02-MAR-2001; 2001US-0272817P.
 05-JUL-2001; 2001US-0303231P.
 12-JUL-2001; 2001US-0305060P.
 10-SEP-2001; 2001US-0318405P.
 12-SEP-2001; 2001US-0318700P.
 04-JAN-2002; 2002US-00037417.
 (CURA-) CURAGEN CORP.
 Kekuda R, Alsobrook JP, Tchernev VT, Liu X, Spytek KA;
 Patturajan M, Grosse WM, Lepnev DM, Burgess CE, Vernet CAM, Li L;
 Gorman L, Edinger S, Sciore P, Ellerman K, Malyankar U;
 Rothenberg M, Stone D, Boldog F, Guo X, Shenoy S, Anderson D;
 Padigaru M, Taupier RJ, Miller CB, Eisen A;
 WPI; 2002-583619/62.
 P-PSDB; ABB09520.
 Novel polypeptides and nucleic acids homologous to transmembrane
 receptor, thymosin, neuromodulin-like family of proteins for diagnosing,
 PT treating cancer, atherosclerosis, neurological, skin and autoimmune
 disorders.
 Claim 9a; Page 121; 323pp; English.
 The invention relates to 24 novel human proteins designated NOV1-NOV14
 (ABB09501-ABB09524), collectively referred to as NOVX proteins, and
 nucleic acids encoding them (ABB093879-ABB093902). NOVX proteins and
 nucleotides are useful in the treatment, diagnosis or prevention of NOVX-
 associated disorders or in the manufacture of a medicament for treating
 such disorders, with specific applications described for each of the 24
 NOVX proteins, based on their homology to known proteins. Various
 disorders are associated with NOVX proteins including neurological
 disorders (e.g., Alzheimer's, Huntington's and Parkinson's diseases),
 pain, behavioural disorders, addiction, tuberosclerosis, cancers
 (e.g., colorectal cancer, leukaemia and osteosarcoma), immune disorders
 (e.g., allergies and autoimmune diseases), myasthenia gravis, asthma,
 various forms of arthritis, diabetes, thyroiditis, cardiovascular disease
 (e.g., hypertension), reproductive disorders, endometriosis,
 incontinence, psoriasis, scleroderma, alopecia, ulcers, pancreatitis,
 cirrhosis, glomerular endotheliosis, polycystic kidney disease, endocrine
 disorders, obesity, bacterial infections and particularly cardiomyopathy,
 atherosclerosis, cell signal processing-related disorders and disorders
 of metabolic pathway regulation. NOVX nucleic acids and polypeptides may
 be used to identify cellular receptors or downstream effectors which
 binds to a NOVX protein, and are also useful as targets for the
 identification of small molecules that modulate or inhibit processes such
 as neurogenesis, cell differentiation, cell motility, cellular
 proliferation, haematopoiesis, wound healing and angiogenesis. NOVX
 nucleic acid sequences can be used to identify a cell or tissue type and
 are useful as a source of primers or probes for forensic biology and for
 identifying and cloning NOVX homologues in other cell types. Cells
 comprising NOVX nucleic acids are useful for producing non-human
 transgenic animals which are useful for studying the function and
 activity of NOVX proteins and for identifying and evaluating modulators
 of NOVX activity. The present sequence represents DNA encoding the
 transmembrane receptor UNC5H2-like protein NOV11. The gene encoding NOV11
 is located on chromosome 10
 Sequence 2895 BP; 557 A; 960 C; 854 G; 524 T; 0 U; 0 Other;
 Query Match 30.1%; Score 866.4; DB 6; Length 2895;
 Best Local Similarity 60.6%; Pred. No. 2.8e-150;
 Matches 1659; Conservative 0; Mismatches 941; Indels 138; Gaps 9;

QY 184 ACCCAGTGCCTGGTGCCAAACCCGAGACCTGTTCCCACTTCTGGTGAGCCGAGGATG 243
 DB 140 ACTCCTTCCCGTCAGCGCCAGCAGAGCGCTGCTCTCTCTCTGAGGAGCCACAGGACG 199
 QY 244 TGTACATCGTCAAGAACAGCCAGTGTCTGTTGTGTGCAAGGCGGTGCGCCGACGACG 303
 DB 200 CTTACTTTCAGTCAACGCGAGTGGGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 259
 QY 304 TCTTCTTCAAGTGCACAAAGGAGTGGGTGCGCCAGGTGGACCACTGATCGAGGCGACGA 363
 DB 260 TCTACTTTCAGTGCACGCGAGTGGGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 319
 QY 364 CAGACGGAGCAGTGTGAGCGGACCATGAGGTCCGATTTAATGTCTCAAGGAGCAGG 423
 DB 320 TGGATGAGGCCACCGGTCTCGGGTGCAGATCGAGGTGTGCGGCGAGCAGG 379
 QY 424 TCGAGAGGTGTTCCGGCTGGAGGAATACTGTGTCAGTGGGTGGCATGAGGCTCTCTGG 483
 DB 380 TGGAGGAGCTCTTTGGGCTGGAGGATTAATGTTGTCAGTGGGTGGGTGGGTGGGTGGGT 439
 QY 484 GCACCAACCAAGAGTCAAGAGGCTTACATCCGATAGCCAGATTGCGCAAGAACTTCGAGC 543
 DB 440 GCACCAACCAAGAGTCCCGAGGCTTACATCCGATAGCCAGATTGCGCAAGAACTTCGATC 499
 QY 544 AGGAGCGCTGGCCAAAGGAGGTGTCCTGGAGCAGGSCATCTGCTGCTCCCTGCGCTCCAC 603
 DB 500 AGGAGCTCTGGGCAAGGAGGTGCCCTGGACCATGAGGTTCTCTGTCAGTGGCGCCGCG 559
 QY 604 CGGAGGSCATCCCTCCAGCGGAGTGGGTCCGAAACGAGGACCTGGTGGAGCCGCT 663
 DB 560 CGGAGGAGGTGCTGTCGCGAGGTGGATGAGTGGATGAGTGGATGATGATGAGTGGATG 619
 QY 664 CCTTGGACCCCAATGTATATACATCAGCGGAGCAGACAGCTGGTGGTGGCAGACGCGCGC 723
 DB 620 CCCAGGACCAACACTTCTGCTCACCATGACCAACACTTATCATCCGCGAGGCGCGCGC 679
 QY 724 TTGCTGACCGGCAACTACACTGCTGGCGCAAGAACATCTGGTGGACGCTGCGCGGAGCG 783
 DB 680 TGTGCGACACTGCAACTATACCTGCTGGCCAAAGAACATCTGGTGGCMAACGCGGAGCA 739
 QY 784 CCTCGCTCTCTCATCGTCTAGCTGAACGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 843
 DB 740 CCACTGCCACCGTCACTGCTAGTGAATGGCGGCTGGTCCAGCTGGGAGAGTGGTCCAC 799
 QY 844 TGTGACGCGCCAGCTGTGGCGCGGCTGTCAGAAACGAGCGGAGCTGCACCAACCGCG 903
 DB 800 CTTGCTCCAAACCGCTGTGCGCGAGGTGCGAGAGCGCACCGGAGCTGCACCAACCGCG 859
 QY 904 CGCTCTCAACGGGGCGCTTTCTGTAGGGGCGAGAAATGTCATGACCGCAGCTCTCTCT 963
 DB 860 CTCCACTCAACGAGGGGCTTCTGCGAGGGCGAGCAATTCAGAA---GACCGCTGCA 916
 QY 964 CTCCTGTTCTCTGTGAGCGACCTGGAGCCGCTGGAGCAAGTGGTGGTGGTGGTGGTGG 1023
 DB 917 CCACCATCTGCCAGTGCATGGGGGCTGGAGGAGTGGAGTGGTGGTGGTGGTGGTGGTGG 976
 QY 1024 TGGACTGCAACCCACTGCGGGAGCGCTGTGAGTGTCTGTGAGGGGCGAGAAATGTCATG 1083
 DB 977 CTGAGTGTGCCACTGGCGTGTGCGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1036
 QY 1084 AGGAGTGCAGGCGCACTGACCTGGACACCCGCACTGTACCACTGACCTGTGTGTACACA 1143
 DB 1037 GTGACTGCAGCGGAGCGCTGTGCTGACTTAAGAACTGCACAGATGGGTGTGCTGCTGCAA 1096
 QY 1144 GTGCTTC-----TGCCCTTGAGGACGTGGGCC 1170
 DB 1097 GTGAGTCAAGTGTGGTCTCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCT 1156
 QY 1171 TCTATGTGGGCTC---ATGCGCGTGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1227
 DB 1157 TGTATGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1216

QY	1228	TCTCGTTTATTTCGCGGAAGAGGAGGGGCTGGACTCAGATGTGGCTGACTGTCATT- 1286
Db	1217	TGTTGGTGTACGCGCGCAACTGCCGTGACTTTCGACACAGACATCACTGACTCATCTGCTG 1276
QY	1287	--CTCAGCTTCAGGCTTCAGCGCGGTGAGCATCAAGCCCGAGCAAGCAGACACCCCGATC 1344
Db	1277	CCCTGACTGTTGTTTCCACCCCGGTCAACTTTAAGACGGCAAGGCCAGTAACCCGCGAGC 1336
QY	1345	TGCT-----CACCATCCAGCCGAGCACTCAGCACCAACCACCTACCAGGCGAGTCTCT 1398
Db	1337	TCCTACACCCCTCTGTGCTCTTAGCTCAGACCGAGCGCGGCATCTACCGGGAACCG 1396
QY	1399	GTCCCGGAGGATGGGCGAGCCCAAGTTCAGCTCACAAATGGGCACTGTCTAGCC 1458
Db	1397	TGTTATGCCCTGAGGACTCCACCGACAAATCCCATGACCAACTCTCTCTGCTGGACC 1456
QY	1459	CC----- 1460
Db	1457	CCTTACCAGCCTTAAGGTCAAGGTCTACAGCTCCAGACCAAGGGCTCTGGGCGAGGCC 1516
QY	1461	-----CTGGGTGGCGCGCCACACACTGTGCACCAACAGCTCTCCCA 1500
Db	1517	TGGCAGATGGGCTGACCTGCTGGGGTCTTGGCGCTGGCACATACCTAGCGATTTCG 1576
QY	1501	CCTGTAGGCGGAGGATTCGTCTCCGCTCTCCACCC-----AGAACTACTTCGGT 1554
Db	1577	CCGGGACACCCACTTCTCTGCACCTGGCAGCGCCAGCCTCGGTTCCAGCAGCTCTTGG 1636
QY	1555	CCCTGCCCGAGGACACGACCAATGACCTTATGGGACCTTCAACTTTCCTGGGGCGCGC 1614
Db	1637	GCCTGCCCGCCAGACCCAGGAGCAGCTCAGCGCACCTTTGGCTGCTGGGTGGGAGGC 1696
QY	1615	TGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCAGATGCCATACCCGAGGGA 1674
Db	1697	TCAGCATCCCGGCACAGGTGTGAGCTTGTGTGTGCCCAATGAGCCATTCGCCAGGCA 1756
QY	1675	AGATCTATGAGATCTACCTCAGCTGCACAAGCGGGAAGACGTGAGTGGCCCTAGCTG 1734
Db	1757	AGTTCTACGAGATGTATCTACTCATCAAGAAGCAGAAAGTACCTCGCCGCTTTCAGAA 1816
QY	1735	GCTGTACAGCCCTGCTGAGTCCCATCGTTAGCTGTGACCCCTGGGTCCTGCTCACCC 1794
Db	1817	GGACCCAGACAGTATTGAGCCCTTCGGTGACCTGTGACCCACAGGCGCTCTGCTGTGC 1876
QY	1795	GGCAGTTCATCCTGGCTATGACCACCTGTGGGAGCCAGCCCTGCAGCTGGAGCCTGC 1854
Db	1877	GCCCGTCACTCTACCAATGCCCACTGTGCCGAGTCAGTGCCCGTGAATGATCTTTC 1936
QY	1855	GCCTCAAAAGCAGTCTGTGCGAGGCACTGGGAGCAGGATGTGTGCACTCTGGCGAGG 1914
Db	1937	AGCTCAAGACCAGGCCCCACAGGCCACTGGG---AGGAGGTGGTGACCCCTGGATGAG 1993
QY	1915	AGGCGCCCTCCACCTCTACTCTGCACAGCTGGAGGCGAGTGCCTGCTAGTCTTTCACCG 1974
Db	1994	AGACCTGAACACACCTCTGCTACTGTCAGCTGGAGCCAGGCGCTGTCACTCTGCTGG 2053
QY	1975	AGCAGCTGGCGCTTTGCCCTGTGGAGAGGCGCCTCAGCGTGGCTGCCGCGCAAGCGCC 2034
Db	2054	ACCAGCTGGGACCTACGTGTTACGGCGAGTCTCTATTCGCGCTCAGCAGTCAAGCGCG 2113
QY	2035	TCAAGCTGCTTGTGTTGGCGGTGTGCTGCACTCTCTCGAGTACAAATTCGGGTCT 2094
Db	2114	TCCAGCTGGCGGTCTTTCGCGCCCGCCCTCTGCACCTCCCTGGAGTACAGGCTCCGGGTCT 2173
QY	2095	ACTGCTGATGACACCAACGATGCACCTCAAGGAGGTGGTGCACCTGGAGAGCAGCTGG 2154
Db	2174	ACTGCTGGAGACACGCGCTGTAGCACTGAAGGAGGTGCTGGAGCTGAGCGGACTCTGG 2233
QY	2155	GGGACAGCTGATCAGAGGACCAAGGCTCTTGCACTTTCAGAGGACAGTTACCAACCTGC 2214
Db	2234	GCGGATACTTGTGTGGAGGAGCCGAACCGCTAATGTTCAAGGACAGTTACCAACCTGC 2293
QY	2215	GCCTATCCATCAGCGATGTGCCACAGCTCCCTGTGGAGAGTAAGCTCTCTCTCAGCTACC 2274

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2004, 12:08:07 ; Search time 197 Seconds
(without alignments)
8115.805 Million cell updates/sec

Title: US-09-970-944-1

Perfect score: 2881

Sequence: 1 agctggggctccggctgag.....ccttccccacacggggaga 2881

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2235.8	77.6	3014	2	US-08-808-982-1
2	2235.8	77.6	3014	3	US-09-306-902A-1
3	1570.4	54.5	1787	2	US-08-808-982-2
4	1570.4	54.5	1787	3	US-09-306-902A-2
5	1147.8	39.8	1282	4	US-09-306-902A-3
6	783.4	27.2	2831	2	US-08-808-982-3
7	783.4	27.2	2831	3	US-09-306-902A-3
8	419	14.5	1605	4	US-09-306-902A-3
9	256	8.9	771	1	US-08-808-982-1
10	119	4.1	305	2	US-08-808-982-4
11	119	4.1	305	3	US-09-306-902A-4
12	51.8	1.8	1280	3	US-08-483-533-38
13	51.8	1.8	1280	4	US-09-283-471A-38
14	51.2	1.8	699	4	US-09-252-991A-7947
15	51.2	1.8	1302	4	US-09-252-991A-7533
16	51.2	1.8	1947	4	US-09-252-991A-7533
17	50.8	1.8	1770	4	US-09-252-991A-12265
18	50.8	1.8	3885	4	US-09-369-364A-16
19	50.8	1.8	4884	4	US-09-252-991A-12126
20	50.8	1.8	4884	4	US-09-252-991A-12292
21	50.2	1.7	1300	3	US-08-483-533-39
22	50.2	1.7	1300	4	US-09-283-471A-39
23	48.6	1.7	1068	4	US-09-252-991A-9933
24	48.6	1.7	1251	4	US-09-252-991A-10142
25	48	1.7	4257	2	US-08-690-473-1
26	48	1.7	4257	3	US-09-259-821A-1
27	48	1.7	4257	3	US-08-843-659-1

Sequence 1048, Ap
Sequence 9, Appli
Sequence 11256, A
Sequence 10995, A
Sequence 11, Appl
Sequence 14, Appl
Sequence 1, Appli
Sequence 3971, Ap
Sequence 4004, Ap
Sequence 32, Appl
Sequence 5, Appli
Sequence 19, Appl
Sequence 1, Appli
Sequence 19, Appl
Sequence 31, Appl
Sequence 6, Appli
Sequence 6, Appli

US-09-620-312D-1048
US-08-440-856A-9
US-09-252-991A-11256
US-09-252-991A-10995
US-09-675-018B-11
US-08-232-463-14
US-08-922-865-1
US-09-510-949-1
US-09-252-991A-3971
US-09-252-991A-4004
US-09-105-537-32
US-09-105-537-5
US-09-320-878-19
US-09-141-908-1
US-09-657-440-19
US-09-340-798A-31
US-07-951-715A-6
US-08-459-448A-6

ALIGNMENTS

RESULT 1
US-08-808-982-1
; Sequence 1, Application US/08808982
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3014 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-808-982-1

Query Match 77.6%; Score 2235.8; DB 2; Length 3014;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;
QY 87 ATGGCGCTCCGGCGGCTGCGCAGCTCTCGGCGCATAGTCTCGCGCTTGGCTC 146
|||||

Db 1 ATGGCGCTCCGGCCCGGCGCTGTGGCAGTGTCTCTGGGCAATAGTCTCTCGCGCCCTGGCTT 60
 QY 147 CCGCGCTCCGGTCCCGCAGCAGATGTCACCGTGGCCAAACCCAGTGCCTGGTGCACACCG 206
 Db 61 CGTGGTTCGGGTCCCGCAGCAGATGTCACCGTGGCCAAATCCAGTCCCGGTGCAACCC 120
 QY 207 GACCTGCTTCCCGACTTCTCGTGGAGCCGAGGATGTGTACATCTCAAGACAGGCCA 266
 Db 121 GACCTGCTCCCGACTTCTCGTGGAGCCGAGGATGTGTACATCTCAAGACAGGCCG 180
 QY 267 GTGCTCTTGTGTGCAAGCCGTCGCGCCAGCAGATCTTCTCAAGTGCACAGCGGAG 326
 Db 181 GTGTTGTGTGTGCAAGCCGTCGCGCCAGCAGATCTTCTCAAGTGCACAGCGGAG 240
 QY 327 TGGGTGGCCGAGGTGACACACGATGATCGAGCGCAGCAGACAGCGGAGCAGTGTGTAGCCG 386
 Db 241 TGGGTGGCCGAGGTGATCAGCTAATGTCAGAGCAGAGTGTGTGTGCGGCTGGAG 446
 QY 387 ACCATGGAGTCCGCAATTAATGTCAGAGCAGAGTGTGTGTGCGGCTGGAG 446
 Db 301 ACCATGGAGTCCGCAATTAATGTCAGAGCAGAGTGTGTGTGCGGCTGGAG 360
 QY 447 GATACGTGGTCCGAGTCCGTCGATGAGTGTCTCGGCGACACCAAGAGTCAGAGGCC 506
 Db 361 GATACGTGGTCCGAGTGTGTGATGAGTGTCTCGGTCACCGAGGATCCCTCCAGCCGAG 420
 QY 507 TACATCCGATAGCCAGATGTCGCAAGAACTTCAGAGCAGAGCGGCTGGCCAGAGGAGT 566
 Db 421 TACATCCGATAGCCAGATGTCGCAAGAACTTCAGAGCAGAGCGGCTGGCCAGAGGAGT 480
 QY 567 TCCCTGAGCAGGATCGTGTCTCGTCCGTCGATGAGTGTCTCGGCGACACCAAGAGTCAGAGGCC 626
 Db 481 TCACTGAGCAAGCAATGTAATCTACCTTGTGCGCCCGCAGAGGAAATCCCGCCAGCTGAG 540
 QY 627 GTGAGTGGTCCGAAAGAGGAGTGTGTGACCGCTCCCTGGAGCCCAATGTATATC 686
 Db 541 GTGAGTGGTCCGAAATGAGGAGTGTGTGACCGCTCCCTGGAGCCCAATGTATATC 600
 QY 687 ACGGGGAGCAGACGCTGT 746
 Db 601 ACGGGGAGCAGACGCTGT 660
 QY 747 TCGTGTGCAAGAACTGT 806
 Db 661 TGTGTGCAAGAACTGT 720
 QY 807 GTGAACGT 866
 Db 721 GTGAACGT 780
 QY 867 GGCTGGCAGAAACGGAGCCGAGT 926
 Db 781 GGCTGGCAGAAACGGAGCCGAGT 840
 QY 927 TGTGAGGGCAGATGT 986
 Db 841 TGTGAGGGCAGATGT 897
 QY 987 AGCTGGAGCCGCTGT 1046
 Db 898 AGCTGGAGTTCGT 957
 QY 1047 CGT 1106
 Db 958 CGGAGT 1017
 QY 1107 GACACCGCACTGT 1166
 Db 1018 GACACCGCACTGT 1077
 QY 1167 GCGCTCTATGTGGGCTGT 1226
 Db 1078 GCTCTCTATCGGCTGT 1137

QY 1227 ATCTCTGTTTATTGCGGAAGAGGAGGCGCTGGACTCAGATGTGGCTGACTGCTGCTT 1286
 Db 1138 GGACTCATTTACTGTGCAAGAGGAGGCGCTGGACTCCGATGTGGCGACTCGTCCATC 1197
 QY 1287 CTCACCTCAGGCTTCCAGCCCTCAGCTCAAGCCAGCAAGCCAGCAAGCCAGCTG 1346
 Db 1198 CTCACCTCGGCTTCCAGCTGTGAGCTCAAGCCAGCAAGCCAGCAAGCCAGCTG 1257
 QY 1347 CTCACCTCAGGCGAGCTCAGCCAGCCAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 1403
 Db 1258 CTCACCTCAGGCGAGCTCAGCCAGCCAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 1317
 QY 1404 CGGAGATGGGCGAGCCAGCCAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 1463
 Db 1318 AGGAGATGGAGCCAGCCAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 1377
 QY 1464 GGTGGCGGCGCACAACATGACCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 1523
 Db 1378 GGGAGTGGCGCCATACGTTGACCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 1437
 QY 1524 TCCGCGCTTCCACCCAGCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 1583
 Db 1438 TCCGCGCTTCCACCCAGCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 1497
 QY 1584 TATGGAGCTTCAACTTCCCTCGGCGCGGCTGATGATCCCTAATACAGGTATACAGCTC 1643
 Db 1498 TACGAGACTTCAACTTCCCTCGGCGCGGCTGATGATCCCTAATACAGGTATACAGCTC 1557
 QY 1644 CTATCCCGCCAGATGCATACCCGAGGGAAGATCTATGAGATCTACCTCAGCTGCAC 1703
 Db 1558 CTATACCCCGGATGCATACCCGAGGGAAGATCTACGAGATCTACCTCAGCTGCAC 1617
 QY 1704 AAGCGAGAGCTGAGTGTGCGCTAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCT 1763
 Db 1618 AAGCGAGAGCTGAGTGTGCGCTAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCT 1677
 QY 1764 AGCTGTGAGACCCCTGCGCTGCTCACCAGGCGAGTCACTCTGCTGCTGCTGCTGCTGCTGCTGCT 1823
 Db 1578 AGCTGTGAGACCCCTGCGCTGCTCACCAGGCGAGTCACTCTGCTGCTGCTGCTGCTGCTGCTGCT 1737
 QY 1824 GGGAGCCAGCCCTGACAGCTGAGCTGCGCTCAAAAGAGCTGCTGAGGAGGAGC 1883
 Db 1738 GGAGAGCCAGCCCTGACAGCTGAGTCTGCGCTCAAAAGAGCTGCTGAGGAGGAGC 1797
 QY 1884 TGGAGCAGAGTGTGCTGACCTGCGGAGGAGCGCTCCACCTCTACTACTGCGAG 1943
 Db 1798 TGGG---AGGATGCTGCACTTGGTGGAGTCACTTCCACCTCTACTACTGCGAG 1854
 QY 1944 CTGAGGCGCAGTGTGCTGCTGCTTCAAGCAGCTGCGGCGCTTGGCTGCTGCTGCTGCTGCTGCTGCT 2003
 Db 1855 CTGAGGCGCGGCTGCTGCTTCAAGCAGCTGCGGCGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1914
 QY 2004 GAGCGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2063
 Db 1915 GAGCGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1974
 QY 2064 TGCACCTCCCTCAGTACCAATCCGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2123
 Db 1975 TGTACGCTCCCTTGT 2034
 QY 2124 AAGAGGT 2183
 Db 2035 AAGAGGT 2094
 QY 2184 CTGACCTCAAGGAGT 2243
 Db 2095 CTGACCTCAAGGAGT 2154
 QY 2244 CTGTGGAAGAGT 2303
 Db 2155 CTGTGGAAGAGT 2214

QY 2430 AACATCAACAGGACACAAGGTTTCTGAGCTGTGGCTCTGGAGAGTGAAGCGGGGTC 2489
DB 1312 AACATCAACAGGACACAAGGTTTCTGAGCTGTGGCTCTGGAGAGTGAAGCGGGGTC 1371
QY 2490 CCAGCCTGTGGGCGGCGGCTTCAAGATCCCTTCCCTCAATTCGGCAGAGATAAT 2549
DB 1372 CCAGCCTGTGGGCGGCGGCTTCAAGATCCCTTCCCTCAATTCGGCAGAGATAAT 1431
QY 2550 TCCAGCCTGGAGCCACACCTGTAGGCGGGGTGCGAGCTGGGAGCTCTGGCCCAAGATC 2609
DB 1432 TCCAGCCTGGAGCCACACCTGTAGGCGGGGTGCGAGCTGGGAGCTCTGGCCCAAGATC 1491
QY 2610 CACCTGGAGCCATCTCAGCTTCTTGGCTTCCAGGCGGCGGCGGCTCTGGCCCAAGATC 2669
DB 1492 CACCTGGAGCCATCTCAGCTTCTTGGCTTCCAGGCGGCGGCGGCTCTGGCCCAAGATC 1551
QY 2670 AACCTGTGGAGCGGCGGCTTCCCAACCGGCAACCTCAGCCAGCTGGCTGACGAGTG 2729
DB 1552 AACCTGTGGAGCGGCGGCTTCCCAACCGGCAACCTCAGCCAGCTGGCTGACGAGTG 1611
QY 2730 GCTGGAGTGGGCGGCGGCTTCCCAACCGGCAACCTCAGCCAGCTGGCTGACGAGTG 2787
DB 1612 GCTGGAGTGGGCGGCGGCTTCCCAACCGGCAACCTCAGCCAGCTGGCTGACGAGTG 1671
QY 2788 GCGGCGGCGGCGGCTTCCCAACCGGCAACCTCAGCCAGCTGGCTGACGAGTG 2845
DB 1672 GCGGCGGCGGCTTCCCAACCGGCAACCTCAGCCAGCTGGCTGACGAGTG 1729
QY 2846 GAAGCGGCGGCGGCTTCCCAACCGGCGGAG 2881
DB 1730 GAAGCGGCGGCGGCTTCCCAACCGGCGGAG 1765

RESULT 5

US-09-833-381-1806
; Sequence 1806, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1806
; LENGTH: 1282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1282)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1806

Query Match 39.8%; Score 1147.8; DB 4; Length 1282;
Best Local Similarity 97.7%; Pred. No. 3.3e-240;
Matches 1259; Conservative 0; Mismatches 18; Indels 11; Gaps 9;
QY 1510 CCGAGGAGTTCCTCCCGGCTCTCCAGCCAGAACTACTTCCGCTCCCTGCCCGGCGGCA 1569
DB 1 CCGAGGAGTTCCTCCCGGCTCTCCAGCCAGAACTACTTCCGCTCCCTGCCCGGCGGCA 60
QY 1570 CCAGCAACATGACCTATGGAGCTTCAACTTCTCGGGGGCGGCTGATGATCCCTAATA 1629
DB 61 CCAGCAACATGACCTATGGAGCTTCAACTTCTCGGGGGCGGCTGATGATCCCTAATA 120
QY 1630 CAGGTATCAGCTCTCTATCCCGGCGGAGTCCATACCGCGGAGGAGATCTATGATCT 1689
DB 121 CAGGAATCAGCTCTCTATCCCGGCGGAGTCCATACCGCGGAGGAGATCTATGATCT 180

QY 1690 ACCTCAGCTTGCACAAGCGGAGAACGAGTTTCCCTAGCTGGCTGTGAGACCTTGC 1749
DB 181 ACCTCAGCTTGCACAAGCGGAGAACGAGTTTCCCTAGCTGGCTGTGAGACCTTGC 240
QY 1750 TGAGTCCCATCTGTTAGCTGTGGACCCCT - GCGTCTCTGCTCACCGGCGAGATCTCT - 1807
DB 241 TGAGTCCCATCTGTTAGCTGTGGACCCCTGGGCGTCTGCTCTCACCGGCGAGTCTCTG 300
QY 1808 GGCTATGAGACCACTGTGGGAGCCAGCCCTGACAGCT - GGAGCTTGGCTCTCAAAAGC 1866
DB 301 GGCTATGAGACCACTGTGGGAGCCAGCCCTGACAGCTGGAGAGCTTGGCTCTCAAAAGC 360
QY 1867 AGTCTGCGAGGCGAGCTGGGAGAGGATGTGCT - TGCACCTGGGCGAGGAGCGCTCTCC 1926
DB 361 AGTCTGCGAGGCGAGCTGGGAGAGGATGTGCT - TGCACCTGGGCGAGGAGCGCTCTCC 418
QY 1927 ACCTCTACTTACTGCGAGCTGAGGCGGAGTGTGCTGCTGCTTCTTACCGAGAGCTGGGC 1986
DB 419 ACCTCTACTTACTGCGAGCTGAGGCGGAGTGTGCTGCTGCTTCTTACCGAGAGCTGGGC 478
QY 1987 GCTTTCCTTGTGGGAGAGGCGCTCAGCGTGGTGGCGCAAGCGCTCAAGCTGCTTC 2046
DB 479 GCTTTCCTTGTGGGAGAGGCGCTCAGCGTGGTGGCGCAAGCGCTCAAGCTGCTTC 538
QY 2047 TGTTCGGGCGGTGGCTTGCCTTCCCTGAGTACAACTCCGGGTCTACTGCTGCTGATG 2106
DB 539 TGTTCGGGCGGTGGCTTGCCTTCCCTGAGTACAACTACTTGTGCTTACTGCTGCTGATG 598
QY 2107 ACACCCAGCATGACATCAAGGAGTGGTGGAGAGCTGGAGAGAGCTGGGGGAGCAGCTGA 2166
DB 599 ACATCAGATGACATCAAGGAGTGGTGGAGAGCTGGAGAGAGCTGGGGGAGCAGCTGA 658
QY 2167 TCCAGGAGCACGGGTCTCTGCACTTCAAGGAGAGTGTACCAACCTGGCTTATCATCC 2226
DB 659 TCCAGGAGCACGGGTCTCTGCACTTCAAGGAGAGTGTACCAACCTGGCTTATCATCC 718
QY 2227 ACGATGTCGCGAGCTCCCTGTGGAGAGTAACTCTTGTGCTGCTACAGGAGATCCCT 2286
DB 719 ACGATGTCGCGAGCTCCCTGTGGAGAGTAACTCTTGTGCTGCTACAGGAGATCCCT 778
QY 2287 TTTATCAGATCTGGAATGGACGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2346
DB 779 TTTATCAGATCTGGAATGGACGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 838
QY 2347 TCAGCCCCAGCACTAGTGAACCTGGCTGCAAGCTGTGGGTGGGAGGAGGCGAGC 2406
DB 839 TCAGCCCCAGCACTAGTGAACCTGGCTGCAAGCTGTGGGTGGGAGGAGGCGAGC 898
QY 2407 GCGAGGCTTCAAGTCAACTTCAAGGAGAGTGTGCTGAGCTGCTGCTGCTGCTGCTGCTG 2466
DB 899 GCGAGGCTTCAAGTCAACTTCAAGGAGAGTGTGCTGAGCTGCTGCTGCTGCTGCTGCTG 958
QY 2467 CTCTGGAGAGTGAAGCGGGTCCAGCGCTGGTGGGCGGAGTCTTCAAGATCCCT 2526
DB 959 CTCTGGAGAGTGAAGCGGGTCCAGCGCTGGTGGGCGGAGTCTTCAAGATCCCT 1018
QY 2527 TCCTCATTTGGCAGAGATAA - TTTCCAGCTTGGACCCAGCTTGGGCGGCTGCTGCTGCTGCTG 2585
DB 1019 TCCTCATTTGGCAGAGATAA - TTTCCAGCTTGGACCCAGCTTGGGCGGCTGCTGCTGCTGCTG 1078
QY 2586 TGGCGG - ACTCTGGCGGAGAACTCCAGCTGGAGAGGCGGCTCTGCTGCTTCTTGGCTTCAA 2644
DB 1079 TGGCGGAACTCTGGCGGAGAACTCCAGCTGGAGAGGCGGCTCTGCTGCTTCTTGGCTTCAA 1138
QY 2645 GCGGAGGCGGAGGCGGCTTCAAGCTGTGGAGGCGGCTCTTCCCGAGCGGCA 2704
DB 1139 GCGGAGGCGGAGGCGGCTTCAAGCTGTGGAGGCGGCTTCCCGAGCGGCA 1198
QY 2705 CCTCAGCGAGCTGGTGCAGAGTGGCTGGAGTGGGCGGAGGCGGCTTCTTCTTCTTCTTCTTCT 2764
DB 1199 CCTCAGCGAGCTGG - TGCAGGAGTGGCTGGAGTGGGCGG - AGCAGAGCTGG - CTCTTCTTCT 1254

QY 2765 ACTGCGGAGGTGAGTGTGAGGCGG 2792
 Db 1255 AGTGTGCGAGGCTGAGTGTGAGGCGG 1282

RESULT 6

US-08-982-3
 ; Sequence 3, Application US/0808982
 ; Patent No. 5939271
 ; GENERAL INFORMATION:
 ; APPLICANT: Tessier-Lavigne, Marc
 ; APPLICANT: Leonardo, E. David
 ; APPLICANT: Hink, Lindsay
 ; APPLICANT: Masu, Masayuki
 ; APPLICANT: Kazuko, Keino-Masu
 ; TITLE OF INVENTION: Netrin Receptors
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 ; STREET: 268 BUSH STREET, SUITE 3200
 ; CITY: SAN FRANCISCO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94104

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA: US/08/808,982
 FILING DATE:
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: OSMAN, RICHARD A.
 REGISTRATION NUMBER: 36,627
 REFERENCE/DOCKET NUMBER: UC96-217
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 343-4341
 TELEFAX: (415) 343-4342
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2831 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cdna
 US-08-982-3

Query Match 27.2%; Score 783.4; DB 2; Length 2831;
 Best Local Similarity 59.4%; Pred. No. 5.7e-161;
 Matches 1625; Conservative 0; Mismatches 971; Indels 139; Gaps 12;

QY 184 ACCAGTGCCTGGTGGCCAAACCGGACCTGCTTCCCACTTCTGCTGGAGCCCGAGGATG 243
 Db 104 ACTCCTTCCCATAGACCCCGGAGCTGCTCCTTCTGCTGAAACGAGGATG 163
 QY 244 TGTATCATCGTCAAGAACCAAGCCAGTGTGCTTGTGTGCAAGGCCGTGCCCGCAGCAGA 303
 Db 164 CCTACATCGTAAAGAACCAAGCCAGTGTGCTTGTGTGCAAGGCCGTGCCCGCAGCAGA 223
 QY 304 TCTTCTCAAGTGCAACCGGGAGTGGTGGCCAGGTGACCACTGATCGAGCCAGCA 363
 Db 224 TCTACTTCAAGTGAATGCGAGTGGGTAGCCAGAAAGGCCACGTACCGCAGGAGGCC 283
 QY 364 CAGACGGGAGCAGTGTGTCAGCGGACCATGAGGTCCGCAATTAATGCTCAAGGCGAGCAGG 423
 Db 284 TGGATGAGGCCACAGGCTTGGCAATACGAGAGTGCATAGAGTGTGCGGCGAGCAGG 343
 QY 424 TCGAGAAGTGTTCGGGTGGAGGAATATCTGGTGCAGTGGCTGGATGGAGCTCTCGG 483
 Db 344 TGGAGGAATTTTGGGGTCGAGGACTACTGGTGTGAGTGGCTGGGCTGGAGCTCTCGG 403

QY 484 GCACACCAAGAGTCAGAAAGGCTCATCCGATAGCCAGATTGGCAAGAACTTCGAGC 543
 Db 404 GAAACCAACAGAGTCGCCGAGCTCATCCGATTCGCTTACCTTGGCAAGAACTTTGACC 463
 QY 544 AGGAGCCGCTGGCCAAAGAGGTGTCCCTGGAGAGGCAATCGTGTGCCCTGCCCTCCAC 603
 Db 464 AGGAGCCCTGTGGCAAGAGGTACCTTGGATCATGAGGTCTCTTGTGAGTGGCCGAC 523
 QY 604 CGGAGGGCATCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTTGGTGGACCGGT 663
 Db 524 CAGAGGGAGTGGCTGTGGCTGAGGTGGAATGGCTCAAGATGAAGATGATCATCATCCG 583
 QY 664 CCTGACACCCCAATGTATATACATCACGCGGGAGCACAGCCCTGGTGGCAGAGCCGCG 723
 Db 584 CTGAGGACACTAATCTCTCTGCTCACCATTGACCAACCTCATCATCCGCGAGCGCGCC 643
 QY 724 TTGCTGACACGGCCAACTACACCTGCTGGTGGCCAAAGAACATCTGTTGGCAGCGCAG 783
 Db 644 TCTCAGACACAGCCAACTACCTGTGTGGCAAGAAATATTGTGGCAAGCGCGCGAGCA 703
 QY 784 CTTCCGCTGCTCATCTGCTACGTGAACGGTGGTGGTGGAGTGGAGCGAGTGGTCCG 843
 Db 704 CGAGGCCACAGTCATCTGTGTGAACGGAGTGGTGGTGGAGTGGAGTGGAGTGGTCCG 763
 QY 844 TCTGACAGCGCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGAGTGTGACCAACCCGG 903
 Db 764 CTGTCTTAACCGCTGGGCGGAGGTGGCAGAAACGCTACTAGGACCTGCAACCAACCCAG 823
 QY 904 CGCTCTTAACGGGCGGCTTTCTGTGAGGGGAGAGATGTCATGACCGGACCGCTCTCT 963
 Db 824 CCCCACTCAATGAGGTGCTTCTGTGAGGAGCAGGCTTGCAGAA---GACGCTTGCA 880
 QY 964 CTCTGTGTGTCTGTGGAGCGGAGCTGGAGCCGCTGGAGCAAGTGGTCCGCTGTGGGC 1023
 Db 881 CCACCGTGTGCCAGTGGATGGAGCGTGGACTGAGTGAGCAAGTGGTCCGCTGTGGGC 940
 QY 1024 TGGACTGCACCCACTGGCGGAGCGGTGAGTGTCTGTGACCCAGCAACCCGCAAGCGAGG 1083
 Db 941 CAGAGTGTGGCACTGGCGAGCGCGAGTGGATGGCAACCGCGCCCGCCAGAACCGAGGCC 1000
 QY 1084 AGGAGTCCAGGGCAGTGAACCTGGACACCGGCAACTGTACAGTGAATCTGTGTACACA 1143
 Db 1001 GTGACTGACGGGACGCTACTTGAATCCAAAGAACTGACCGGATGGGCTGTGGCTGTGA 1060
 QY 1144 GTGCTTCTGGCC-----CTGAGGACCTGGGCC 1170
 Db 1061 ATCAGAGAACTCTTAAACGACCCCTTAAAGCCGCCCTTGGAGCCGTGGGAGACGTGGCG 1120
 QY 1171 TCTATGTGGGCTC---ATCGCGTGGCCGCTGTGCCCTGGTCTCTGCTGTCTTGTCTCA 1227
 Db 1121 TGTATGGGCGCTCGTGTGGCCGCTCTTGTGTGTTCTGGCAGTCTCTCATGCTGTAGGAG 1180
 QY 1228 TCCTGTTTATTGCGGAAAGAGGAGGGGTGGAGTGGATGATGCTGCTGCTGCTGCTCAAT- 1286
 Db 1181 TGATCGTGTACCGGAGAACTGCCGGGACTTCGACACGGACATCACTGACTCTCTCTGTG 1240
 QY 1287 ---CTCAGCTCAGGCTTCCAGCCCGCTCAGACCTCAAGCCCGCAGCAAGCAACCCCATC 1344
 Db 1241 CCCTCAGTGTGGTTCACCCCGCTCACTTCAAGCTGCAAGCCCGCCAGCAACCCAGCAGC 1300
 QY 1345 TGCT-----CACCATCCAGCCGAGCTTCAGCAACA---CCACCCTTACAGGGCAGTC 1395
 Db 1301 TCCTGACCCCATCCGCCCTCCGACCTTAACGGCCAGTGTGGCATCTTACCGCGAGACCTG 1360
 QY 1396 TCTGTCCCGGACGA----- 1411
 Db 1361 TGTATGCCCTCAGGACTCTGCCGCAAGATCCCTATGACTAATTCACCCCTTCTGGATC 1420
 QY 1412 -TGGGCCAGCCCAAGTTCAGCTCACCAGTGGGACCTGCTCAGCCC-----CC 1461
 Db 1421 CTTGCCCGAGCTCAAGATCAAGTCTATGACTCCAGCACCACCTCGCTCTGGGGCTGGCC 1480
 QY 1462 TGGGTGGCGCGCCGACACACATGACACACAGCTCTCCCACTCTGAGGCCGAGGATTCG 1521

1481 TGGCTGATGGAGCCGACCTGCTGGGTGTTTACCAACCCGGTACATACCAGCGATTCTT 1540
1522 TCTCCCGCTCTCCACCCAGAAC-----TACTCCGCT 1554
1541 CCCGGACACCACTCTCTGCACCTGCGAGCGCCAGCCCTTGGTCCACGACCTCTCTGG 1600
1555 CCTGCCCCGAGGACCCAGCAACATGACCTATGGACCTTCAACTTCTCGGGGCCCGC 1614
1601 GCCTCCCTCGAGACCCAGCAGCAGTGTGAGTGGCACCTTTGGTTGCCCTGGGTGGAGGC 1660
1615 TGATGATCCCTAATAFACGATATCAGCTTCTCATCCCCCAGATGCCATACCCGAGGGA 1674
1661 TGACCAATCCCGGACACAGGGGTGAGCTTGGTACCAAAATGGAGCCATTCGCCAGGCA 1720
1675 AGATCTATGAGATCTACCTCAGCTGACACAGCGGAGAGAGTGAAGTGGCCCTAGCTG 1734
1721 AGTTCTATGACTTGTATCTAGTATCAACAGACTGAAGACCCCTCCCATCTTGGGAAG 1780
1735 GCTGTGAGACCTGCTGAGTCCCATGTTAGTGTGGAACCCCTGGCGCTCTGCTCAACC 1794
1781 GTTCCACAGAGTATGAGCCCTCGGTGACCTGGGGCCCAAGGCTCTCTCTGTGCC 1840
1795 GGCCAGTCACTCGCTATGACACTGTGGGAGCCCGAGCCCTGACAGCTGAGCCCTGC 1854
1841 GCCCTGTTGCTCAGTGTGCCCTGCTGAGTCAATTGCCGGAGACTGGATCTTCC 1900
1855 GCCTCAAAAGCAGTCTGTGGAGGCGAGCTGGAGCAGGAGTGTCTGCACCTGGCGAGG 1914
1901 AGCTCAAGACCCAGGCCCATCAGGCCCACTGGG---AGGAGTGGTGACTTTGGATGAG 1957
1915 AGGGCCCTCCGACCTTACTACTGTCAGCTGGAGCCAGTGCCTGCTAGTCTTTCACCG 1974
1958 AGACTGTGAACCCCTGCTACTGCCAGCTAGAGGCTAAATCTGCCACATCTGTGTG 2017
1975 AGCAGCTGGCCGCTTGGCCCTGGTGGAGAGGCCCTAGCGTGGCTGCGGCCAAGGCC 2034
2018 ACCAGCTGGGTACTAGTGTTCAGGGGCGAGTCTACTCCCGCTCCGAGTCAAGCGC 2077
2035 TCAAGCTGCTTCTGTTGCGCCGGTGGCTGCACTCCCTCGAGTCAACATCGGGTCT 2094
2078 TCCAGCTAGCACTTCGCCCCAGCCCTCTGCACCTCCCTGGAGTATAGTCTCAGGGTCT 2137
2095 ACTGCTGTGATGACACCCAGATGCACTCAAGAGGTGTGCACTGGAGGAGCAAGCAGCTG 2154
2138 ACTGTGGAGGACACTCCTGACGACTGAAGAGGTCTAGAGCTGGAGAGACTCTGG 2197
2155 GGGGACAGCTGATCCAGGAGCCAGGGTCTGCACTTCAAGGACAGTTACCAACCTGC 2214
2198 GTGGCTACTTGTGGAGGAGCCCAAGACTTGTCTTTAAGGACAGTTACCAACCTAC 2257
2215 GCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTGAAGTCTGTGACGTAC 2274
2258 G-CTCTCCCTCCATGACATCCCTCCATGCCCACTGGAGGAGCAAACTACTGGCAAGTACC 2316
2275 AGGAGATCCCTTTATACATCTGGAATGGCAGCAGCGGTACTTGCACCTGCACCTTCA 2334
2317 AGGAGATCCCTTCTACCATGTGTGAAGCGCAGCAGAAAGCCCTGCACTGCATTTCA 2376
2335 CCTGGAGGCTCAGCCCAAGCACTAGTGAAGCTGGCTGCAAGCTGTGGGTGTGGCAGG 2394
2377 CCTGGAGGACATAGCCCTAGCTCCACTGAGTTCACTGTAAAGTCTCGTGGCGCAGG 2436
2395 TGGAGGGCAGCGGAGAGCTTTGAGCATCACTTCAACATCAC---CAGGACACAGGT 2451
2437 TAGAAGGGAAGCCAGATTTTCCAGTGCACACACGCTGCTGAGAGCCTGTGCTGGCT 2496
2452 TTGCTGAGTGTGCTGTGAGAGTGAAGCGGGGTCCAGCCCTGTGGGCCCCAGTG 2511
2497 CCTGGATGCACTCTGCTCTGCCCTGGCAATGTCTGCCACACACAGCTGGGACCTATG 2556
2512 CTTTCAAGATCCCTTCTCTATTCGGGAGAGATTAATTTCCAGCTGGAACCCACCTGTA 2571

Db 2557 CTTTCAAGATACCAGTGTCCATCCGCCAAGAAAGATCTGCAACAGCCTGGAAGCCCAACT 2616
QY 2572 GCGGGGTGCGGACTGGGGGACTCTGGCCCAAAACTCCACCTGGACAGCCATCTCAGCT 2631
Db 2617 CACGGGGCAATGACTGCTGGCTGTTGGCAGCAAGAGCTCTCCATGAGCCGGTACCTGAAT 2676
QY 2632 TCTTTGCTTCCAGCCAGCCAGCCCAAGCCATGATCCTCAACCTGTGGAGGCGCGCACT 2691
Db 2677 ACTTCGCCACCAAAAGTGTGCTCCACAGGCTGATCTTAGACCTCTGGGAAGCTCGGCAGC 2736
QY 2692 TCCCCAAGCGCAACTCAGCCAGCTGGCTGAGCAGTGGCTGGAGTGGCCAGCCAGCAGG 2751
Db 2737 AGGATGATGGGACCTCAACAGCCTGGCCAGTGCCTTGGAGGAGATGGGCAAGAGTGAGA 2796
QY 2752 CTGGCTCTTCCAGAGTGTGGAGGCTGAGTGTCTGA 2786
Db 2797 TGCTGTAGCCATGACCACTGATGGCGATTGCTGA 2831

RESULT 7

US-09-306-902A-3
; Sequence 3, Application US/09306902A
; Patent No. 6277585
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsey
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/306,902A
; FILING DATE: 07-May-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2831 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-306-902A-3

Query Match 27.2%; Score 783.4; DB 3; Length 2831;
Best Local Similarity 59.4%; Pred. No. 5.7e-161;
Matches 1625; Conservative 0; Mismatches 971; Indels 139; Gaps 12;
QY 184 ACCAGTGCCTGGTGCACACCGGAGCCTGCTTCCCACTTCTGCTGGAGCCCGAGGATG 243
Db 104 ACTCCTTCCCATCAGACCCCGGAGAGCTGCTCCTTCTGCTGGAACGAGAGATG 163
QY 244 TGTATATGTCAGAGAACAGCAGTGTGCTTGTGTGCAAGGCGGTGCTCCGCCAGCAGA 303

164 CCTACATCGTAAGAACACAGCCAGTGGAAATGCCACTGCGAGCCCTTCCCTGCCACACAGA 223
 304 TCTTCTTCAAGTGCACACGGGGAGTGGGTGGCGCAGGTGACACACGCTGATCGAGCGCAGCA 363
 224 TCTACTTCAAGTGTATGGCGAGTGGGTAGCCAGAAAGCCACGTCACGAGGAGAGCC 283
 364 CAGACGGGAGCAGTGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGG 423
 284 TGGATGAGGCCACAGGCTTGGCAATACGAGAGGTGCAGATAGAGGTGTGCGCGCAGCAGG 343
 424 TCGAGAAAGTGTTCGGGCTGGAGGAATACGTGGTCCAGTCCGTCGTCGATGAGTCTCTCGG 483
 344 TGGAGGAACTTTTGGGCTCGAGGACTACTGGTGTGAGTGGTGGCTGGAGTCTTCGG 403
 484 GCACACCAAGAGTCAAGAGGCTTACATCCGATAGCCAGATTTGGCGCAAGAACTTCGAGC 543
 404 GAACCAACCAAGTTCGCGAGCCTACATCCGCATTGCCTACTTTGGCGCAAGAACTTTGACC 463
 544 AGGAGCGCTGGCCAAAGGAGGTGTCCCTGAGCAGGCGCATCGTGTGCTGCCCTGCGGTCCAC 603
 464 AGGAGCCTCTGGCGAAGAGGTACCCCTTGGATCATGAGGTCTTCTGCACTGCGGCCAC 523
 604 CGAGGGCATCCCTCCAGCCGAGGTGAGTGGTCCGGAACGAGGACCTGTGTGACCCGT 663
 524 CAGAGGGAGTGCCTGTGGCTGAGGTGGAATGGCTCAAGAAATGAAGATGTCTATGATCCCG 583
 664 CCTGGACCCCAATGTATACATACCGCGGAGCAGCCTGGTGGTGGCAGCGCCCGCC 723
 584 CTAGGAACACTAATCTCTGCTACCATGACCATGACACACCTCATCATCCGCGAGCGGCC 643
 724 TTGTGTACAGCGCCCAACTACACTGCGTGGGCAAGAACATCGTGGGACGTCGCGCGAGCG 783
 644 TCTCAGACACAGCCAACTACACCTGTGTGGCAAGAAATATTGTGGCCCAAGCGCGGAGCA 703
 784 CTTCCGTGTGTGTATGTCTACGTGAACGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 843
 704 CGACGGCCACAGTATCGTCTATGTGAACGAGGTGTGTGACGTGGGCAAGATGTGTAC 763
 844 TCTCAGCGCCAGCTGTGGCGCGCTGGCAGAAACGAGCGGAGCTGCAACCAACCGG 903
 764 CTTGCTCTAACCGCTGCGCGCGAGGTGGCAGAAACGTAAGTACCTGCAACCAACCGAG 823
 904 CGCCTCTCAACCGGGCGCTTTCTGTGAGGGCGAGAAATGTCCATGACCGCACCGTCTCT 963
 824 CCCCACTCAATGGAGTGCCTTCTGCGAGGACAGGCTTGGCAGAA---GACGCTTGCA 880
 964 CTCTGCTGTCTCTGTGACGCGCAGCTGGAGCCGCTGGAGCAAGTGTGCGCTGTGGGC 1023
 881 CCACCGTGTGCGCCAGTGGATGGAGCGTGGACTGTGAGTGGAGCAAGTGTGCGCTGACGA 940
 1024 TGGACTGCACCCACTGGCGGAGCCCTGAGTGTCTGACCCAGCACCCCGCAACGAGGGG 1083
 941 CAGAGTGTGCGCACTTGGCGAGCCGAGTGTGATGGCAGCCGCGCCCGCAACGAGGCC 1000
 1084 AGGAGTGCAGGGCACTGACCTGGACACCCGCAACTGTACAGTGAACCTCTGTGTACACA 1143
 1001 GTGACTGCAGCGGACGCTTACTTGAATCCAGAACTGCACCGATGGGCTGTGCGTGA 1060
 1144 GTGCTTCTGGCC-----CTGAGGACGTGGGCC 1170
 1061 ATCAGAGAACTTAAACGACCCATAAAGCGCCCTCGAGCGCTGCGGAGAGCTGGCGC 1120
 1171 TCTATGTGGGCTC---ATCGCGGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1227
 1121 TGTATGGCGGCTCGTGTGGCGCTTTTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1180
 1228 TCTCTTTTATTGCGGGAAGAGGAGGGCTGACCTCAGATGTGCTGCTGCTGCTGCTGCTGCT 1286
 1181 TGTATGTGTACCGGAGAACTGCGGGACTTGCACGAGACATCACTGACTCTCTGTGCTG 1240
 1287 --CTACCTCAGGCTTCCAGCCGCTCAGCATCAAGCCCGCAGCAAGCAACACCCCATC 1344

1241 CCTCACTGGTGGTTCACCCCGTCAACTTCAAGACTGCAAGGCCAGCAACCCACAGC 1300
 1345 TGCT-----CACCATCCAGCCGACCTCAGCACCA---CCACCACCTTACCAGGGCAGTTC 1395
 1301 TCTGTACCCATCCCGCCCTCCGACCTTAAACGCGCAGTGTCTGGCATCTTACCGCGGACCTG 1360
 1396 TCTGTCCCCCGCAGGA----- 1411
 1361 TGTATGCCCTTCAGGACTCTGCCGACAAGATCCCTATGACTAATTCACCCCTTCTGATC 1420
 1412 -TGGGCCACAGCCCAAGTTCAGAGTCAACATGGGCACTGTCTCAGCCC-----CC 1461
 1421 CTTTGGCCAGCCTCAAGATCAAGTCTATGACTCCAGACCATCGGCTCTGGGGTGTGGCC 1480
 1462 TGGGTGGCGCGCCACACACATGACAGCTCTCCACCTCTGAGGCGCGAGAGTTCG 1521
 1481 TGGCTGATGAGAGCGGACCTGTGTGGTGTCTTACCACCCGGTACATACCCAGGCGATTTCT 1540
 1522 TCTCCCGCCTCTCCACCCAGAAC-----TACTTCGGCT 1554
 1541 CCGGGACACCCACTTCTCTGCACTGGCGAGCGCCAGCCTTGGTCCCAGCACTCTCTGG 1600
 1555 CCTTGGCCCGAGGACACAGCAACATGACCTATGGGACCTTCAACTTCTCGGGGCGCGGC 1614
 1601 GCCTCCCTCGAGACCCCGAGCAGTGTCACTGGCACCTTTGGTGTGCTGTGGTGGAGGC 1660
 1615 TGATGATCCCTAATACAGGTATCAGCTCTCTCATCCCGCCAGATGCCATACCCCGAGGGA 1674
 1661 TGACCATTTCCGGGACAGGGGTGAGCTGTGTGTAACCAATGAGGCCATTTCCCGAGGCA 1720
 1675 AGATCTATGAGATCTACTCTCACGCTGCACAAAGCGGAGACGAGTGGTGGCTTCTAGCTG 1734
 1721 AGTTCTATGACTTGTATCTAGTATCAACAAGACTGAAGCACCCCTCCCACTTTTCGGAAG 1780
 1735 GCTGTACAGCCCTGCTGAGTCCCATCGTGTAGCTGTGACCCCTCGGCTGTCTGTCACCC 1794
 1781 GTTCCACAGACGATTTAGAGCCCTCGGTGACCTCGGGGCCACCGGGCTCTCTCTGTGCC 1840
 1795 GGCCAGTCACTCTGGCTATGGACCACTGTGGGAGCCCGACCTCTGACAGCTGGAGCCCTGC 1854
 1841 GCCCTGTGTCTCTACCTGTGCCCACTGTGTGAAGTCAATTGGCCGAGACTGGATCTTCC 1900
 1855 GCCTCAAAAAAGCAGTCTGTGGAGGCGAGCTGGGAGCAGGAGTGTGTGACCTGGCGAGG 1914
 1901 AGCTCAAGACCCAGGCCCATCAGGGCCACTGGG---AGGAGGTGGTGAATTTGGATGAGG 1957
 1915 AGGCGCCCTCCCACTCTACTTACTGACAGTGGAGGCGAGTGCCTGTCTACCTTCTCACCG 1974
 1958 AGACTCTGAACACACCCCTGTCTACTGCCAGCTAGAGGCTTAAATCTCTGCCACATCTCTGTGG 2017
 1975 AGCAGCTGGGCGCTTTTCCCTGTGGGAGAGGCGCTCAGCGTGGCTGCCCGCAAGCGCC 2034
 2018 ACCAGTGGGTACCTACGTGTTCAGGGCGAGTCTCTACTCCCGCTCCGAGTCAAGCGGC 2077
 2035 TCAAGCTGTCTTCTGTGTGGCGCGTGTGCACTCCCTCGAGTACAACTCCGGTCT 2094
 2078 TCCAGTAGCCATCTTGGCCCGACGCTCTGCACCTCCCTGGAGTATAGTCTCAGGGTCT 2137
 2095 ACTGCTGCTATGACACCCAGCATCAAGAGGTGTGAGCTGAGAGCAGAGCTGG 2154
 2138 ACTGTCTGGAGGACACTCTCGAGCACTGAAGAGGTCTTAGAGTGGAGGAGTCTTGG 2197
 2155 GGGGACAGCTGATCCAGGAGCCACCGGTCCTGCATCTCAAGGACAGTTTACCAACCTGCG 2214
 2198 GTGGGTACTTGTGGAGGAGCCCAAGACTTTGCTCTTTAAGGACAGTTTACCACAACCTAC 2257
 2215 GCCTATCATCAACGATGTGCCAGCTCTGTGTGAAGAGTAACTCTTGTGCTGAGTACC 2274
 2258 G-CTCTCCCTCATGACATCCCGCATGCCCCACTTGGAGGAGCAAACTTACTGGCCAACTAC 2316
 2275 AGGAGATCCCTTTTATCAGATCTGGAATGGACCGAGCGGTACTTGTGACTGACCTGCA 2334
 2317 AGGAGATTCCTTCTACCATGTGTGGACGCGAGCCAGCAAGAGCCCTGCACCTGCTTCA 2376

```
QY 2335 CCCTGGAGCGGTGTCTAGCCCGCAGCACTAGTGAACCTGGCCCTGCAAGCTGTGGTGTGGCAGG 2394
D 2377 CCCTGGAGAGACATAGCCTTAGCCTCACTGATGTTACCTGTAAAGTCTGCTGGCGCAGG 2436
QY 2395 TGGAGGCGGAGCGGAGAGCTTTCAGCATCACTTCAACATCAC---CAAGGACACAAGGT 2451
D 2437 TAGAAGGGGAAGGCGCAGATTTTCCAGCTGCACACACAGCTGCCTGAGACGCTGCTGCTGCT 2496
QY 2452 TTGCTGAGCTGCTGGCTCTGAGAGTGAAGCGGGGTCCAGCCCTGTGTGGGCCCGCAGTG 2511
D 2497 CCCTGGATGACTCTGCTCTGCCCTTGGCAATGCTGCGCCACACACAGCTGGGACCCCTATG 2556
QY 2512 CTTTCAAGATCCCTTCTCATTTGGCAGAGATAAATTTCCAGCTTGGACCCACCCCTGTA 2571
D 2557 CTTTCAAGATACCACTGTCCATCCGCCAGAGATCTCAACAGCTGGACGCCCCCAACT 2616
QY 2572 GCGGGGTGCGGACTGGCGGACTCTGGCCGAGAACTCCACTGGACAGCCATCTAGCT 2631
D 2617 CACGGGCAATGACTGGCGGCTGTTGGCAGAGAGCTCTCCATGGACCGGTACCTGAAT 2676
QY 2632 TCTTTTGCTCCAAAGCCAGCCCGCAGCCAGCATGATCCTCAACTGTGGAGCGGGCACT 2691
D 2677 ACTTGGCCACCAAGCTAGTTCACAGCGGTGATCTTAGACCTCTGGAGAGATGGGCAAGTGAGA 2736
QY 2692 TCCCCAACGGCAACCTCAGCCAGCTGGCTGCGCAGTGGCTGGAGTGGCTGGCCAGCCAGCG 2751
D 2737 AGGATGATGGGACCTCAACAGCCTGCCAGTGCCTTGGAGGAGATGGGCAAGTGAGA 2796
QY 2752 CTGGCTCTTTCAGCTGTCGAGGCTGAGTGCTGA 2786
D 2797 TGCTGTAGCCATGACCACTGATGGCGATTGCTGA 2831

RESULT 8
US-09-833-381-1807
; Sequence 1807, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1807
; LENGTH: 1605
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(1605)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1807

Query Match 14.5%; Score 419; DB 4; Length 1605;
Best Local Similarity 62.4%; Pred. No. 6.6e-82;
Matches 737; Conservative 0; Mismatches 435; Indels 10; Gaps 5;

QY 1556 CTGCCCCGAGGACACAGCAATGACCTATGGAGCTTCACTTCTCTCGGGGCGCGCT 1615
D 421 CTTGCCCCGAGACCCAGGAGCAGCGTTCAGCGCACCTTTGGCTGCTGGGTGGAGGCT 480
QY 1616 GATGATCCCTAATACAGTATCAGCTCTCTATCCCGCCAGATGCCATACCCCGAGGAA 1675
D 481 CACATCCCGGACACAGGCGTTCAGCTTGTGTGTCCTTCCATGAGCCATTCACAGGGCA 540
QY 1676 GATCTATGA--GATCTACTCAGCTGCACAGCGGAGAGCTGAGGTGCCCCTAGCT 1733
D 541 GTTCTACGAAGATGATTCTACTCATCAACAGGAGAAAGTACCTCCCGCTTTCAGAA 600
```

RESULT 9

US-08-253-155A-17/c

; Sequence 17, Application US/08253155A

```
QY 1734 GGCTGTCTAGACCCCTGTGAGTCCCATCGTTAGCTGTGGACCCCTGGCGTCTCTGCTACC 1793
D 601 GGGACCCANACAGTATTGAGCCCTCGGTGACCTGTGGACCCACAGGCTCTCTGCTGTC 660
QY 1794 CGGCGAGTCACTCTGGCTATGACACACTGTGGGAGCCAGCCCTGACAGCTGGAGCCTG 1853
D 561 GCGCCCGTCACTCTCAACATGCCCACTGTGCCAAGTCAGTGCCCGTGAATGATCTTT 720
QY 1854 GCGCTCAAAAGCAGTCTGTGCGAGGCGAGCTGGGAGCAGGATGTGCTGACCTGGCGAG 1913
D 721 CAGCTCAAGACCCAGGCGCCACAGGCGCACTGGG---ANGAGGTGGTACCCCTGATAG 777
QY 1914 GAGCGCCCTCCCACTCTACTACTGCGAGCTGGAGGCGAGTGTGCTGCTTCACTTACC 1973
D 778 GAGACCTTGAACACACACCTGCTACTGCGAGCTGGAGCCAGGCGCTTGCATCTCTGCT 837
QY 1974 GAGCAGCTGGGCGCT--TTGCCCTGTGGAGAGGCGCTCAGCTGGGTGGCGCAAGCG 2032
D 838 GACAGCTGGGCGCTTACCGGTGTTACCGGGCGAGTCTTATTCGCGCTCAGCAGTCAAGCG 897
QY 2033 CCTCAAGCTGCTTCTGTGCGCGGTGGCTGACCTCTGCGAGTACACATCCCGGT 2092
D 898 GCTCCAGCTGCGCGTCTTTCGCGCGCGCTTGCACCTCTCTGGAGTACAGCTTCCCGGT 957
QY 2093 CTACTGCTGCATGACACACCCAGCATGCACTCAAGGAGTGTGCGAGTGGAGAGCAGCT 2152
D 958 CTACTGCTGGAGGACACGCTGTAGCAGTGAAGGAGTGTCTGGAGCTGGAGCGGACTCT 2107
QY 2153 GGGGGGACAGCTGATCCAGGAGCCAGGCTCTGCACTTCAAGGACAGTGTACCAACCT 2212
D 1018 GGGCGGATCTTGTGGAGGAGCGGAAACCGCTTAATGTTCAAGGACAGTGTACCAACCT 2077
QY 2213 GCGCTATCCATCCAGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTCCTCTGCTAGCTA 2272
D 1078 GCGCTCTCCCTCCATGACCTCCCGCATGCCATTTGGAGGAGCAGCTGTGCCAAATA 1137
QY 2273 CAGGAGATCCCTTTTATCATCTGGAATGCCAGCGAGCGGTACTTGCATCTGCACCT 2332
D 1138 CCAGGAGATCCCTTTTATCATCTGGAATGCCAGCGAGCGGTACTTGCATCTGCACCT 1197
QY 2333 CACCTTGGAGCGGTGTCAGCGCCAG--CACTAGTACCTGGCTGCAAGCTGTGGGTGGC 2391
D 1198 CACCTTGGAGCGGACAGCGCTTGGCTTCCACAGAGCTCACCTGCAAGATCTGCGTGGCG 1257
QY 2392 AGGTGGAGGCGGACGCGGAGCTTTCAGCATCAACTTCAACATCACCAAG--GACACAA 2448
D 1258 AAGTGGAGGCGGAGGCGCAGATATTCAGCTGCATACCACTCTGGCAGAGACACCTGCTG 1317
QY 2449 GGTGTGCTGAGCTGTGGCTCTGGAGAGTGAAGCGGGGTGCCAGCTTGGTGGGCCCCA 2508
D 1318 GCTCCCTGGACACTCTCTGCTCTGCGCTGGCAGCACTGTCAACCCAGCTGGGACCTT 1377
QY 2509 GTGCTTCAAGATCCCTTCTCATTCGCGCAGAGATAATTTCCAGCTGGAGCCCACT 2568
D 1378 ATGCTTCAAGATCCCACTGCTCCATCCCGCCAGAGATATGCAACAGCTAGATGCCCA 1437
QY 2569 GTAGGCGGGGTGCCAGCTGGCGGAGCTCTGGCCACAGAACTCCACCTGGACAGCATCTCA 2628
D 1438 ACTCACGGGCAATGACTGGCGGATGTAGCAGCAGAGCTCTCTATGACCCGTTACCTGA 1497
QY 2629 GCTTCTTTGCTTCAAGCCCGAGCCCGCAGCCATGATCTCTCAACTGTGGAGGCGCGC 2688
D 1498 ATTAATTTTCCCAACAAAGCGAGCGCCCGCTGATCTCTGACCTCTGGAGTCTGTC 1557
QY 2689 ACTTCCCCAAAGGCAACCTCAGCCAGCTGGCTGCGAGAGTGG 2730
D 1558 AGCAGGACGAGGAGACCTCAACAGCTGCGAGTGGCTTGG 1599
```

Thu Oct 7 10:58:54 2004

```

; Patent No. 5691147
; GENERAL INFORMATION:
; APPLICANT: Gyuris, Jenio
; APPLICANT: Draetta, Giulio
; TITLE OF INVENTION: CDK4 Binding Proteins
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; FILING DATE: 02-JUN-1994
; APPLICATION NUMBER: US/08/253,155A
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: M11-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-253-155A-17

Query Match      8.9%; Score 256; DB 1; Length 771;
Best Local Similarity 67.2%; Pred. No. 1.3e-46;
Matches 406; Conservative 1; Mismatches 191; Indels 6; Gaps 3;

QY 1741 AGACCTGCTGAGTCCCATGTTAGCTGTGACCCCTGGGCTCTGCTACCCGGGCGAG 1800
DB 738 AGACAGTATTAGAGCCCTCGGTGACCTGTGACCCAGAGGCTCTGCTGTGCGCCCGG 679
QY 1801 TCATCTCTGGCTATGGACCACTGTGGGAGCCAGCCCTGACAGCTGGAGCCTGGCGCTCA 1860
DB 678 TCATCTCTACCATGCCCACTGTGCCGAGTCACTGCGGCTGCTGCTGCTGCTGCTGCT 619
QY 1861 AAAAGCAGTCTGCGAGGCGAGCTGGGAGCAGGATGTGCTGCTGCTGCTGCTGCTGCTG 1920
DB 618 AGACCCAGGCCACCAGGCGCACTGGG---AGGAGGTGCTGACCTCTGGATGAGGAGCCC 562
QY 1921 CCTCCCACTCTACTACGCTGAGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
DB 561 TGAACACACCTGCTACTGSCAGCTGGAGCCAGGCGCTGTCACTGCTGCTGCTGCTGCTG 502
QY 1981 TGGGCGGCTTTGCCCTGTGGGAGAGGCGCTCAGCGTGGCTGCGCCCAAGCGCTCAAGC 2040
DB 501 TGGGCACTACGTGTTTACGGGCGAGTCTTATCCCGCTCAGAGTCAAGCGGCTCCAGC 442
QY 2041 TGCTTCTGTTTGGCGCGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
DB 441 TGGCGGTCTTGCGCCCGCCCTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 382
QY 2101 TGCATGACCCACGATGACCTCAAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
DB 381 TGGAGGACAGGCTGTAGCACTGAAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 322
QY 2161 AGCTGATCAGGAGCCAGCGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
DB 321 ACTTGTGGAGGAGCGGAACCGCTAAATGTTTCAAGGACAGTGTAAACACAACTTGGC-CTCT 263

```

```

QY 2221 CCATCCAGATGTGCCAGCTCCCTGTGGAAGTAAGTCTCTTGTCTAGTACCGAGGAGA 2280
DB 262 TCCTCATGACTCTCCCATGCCCATTGGAGGAGCAAGCTGCTGGCCAAATACCGAGGAGA 203
QY 2281 TCCCCTTTTATCAGATCTGGAATGGCACGCGGCTATTGCACTGCACTTCCACCTGG 2340
DB 202 TCCCCTTCTATCAGATTTGGAGTGGAGCCAGAGG--CCTCCACTGCACTTTCCACCTGG 145
QY 2341 AGCG 2344
DB 144 AGAG 141

RESULT 10
US-08-808-982-4
; Sequence 4, Application US/080808982
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 305 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-808-982-4

Query Match      4.1%; Score 119; DB 2; Length 305;
Best Local Similarity 68.1%; Pred. No. 5.6e-17;
Matches 209; Conservative 0; Mismatches 95; Indels 3; Gaps 3;

QY 1906 TGGCGGAGGAGCGCCCTCCACCTCTACTACTGCACTGGAGGCGGAGGCGCTGAGCGTGG 1965
DB 1 TGGATGAGGAGACCTGNACACACCTGCTACTG-CAGCTGAGGAGCGGCGCTG-TACA 58
QY 1966 TCTTACCGAGCAGTGGGCGCTTTGCTGCTGGGAGAGGCGCTGAGCGTGGTGGCG 2025
DB 59 TCTGCTGGACAGCTGGGCACTACGTTTTCACGGGCGAGTCTCTATTTCCCGCTCAGCAG 118
QY 2026 CCAAGCGCTCAAGCTGCTTCTGTTGGCGCGGTGGCTGCACTTCCCTCGAGTACAACA 2085
DB 119 TCAAGCGGCTCCAGCTGGCGGT-TTCGCCCCCGCCCTCTGCACTTCCCTGGAGTACAGCC 177

```


Db 615 ACCCCGCGACCCCGCGACCCCGCGACCCCGCGCGGCTGCGCTTCTGCCCCACGTC 674
QY 126 ATAGTCTCGCGCTTGGCTCGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAAAC 185
Db 675 CGGGTGGCGCACTGGTGGTCTGGGCTCGGCGCCCGGCTGGCGCGCGGCTCGTGG 734
QY 186 CCAGTGCCTGGTGCCCAACCGGACCTGCTTCCCACTTCTGCTGGAGCCCGAGGATGT 244
Db 735 GCCCGGAGCGGGCCGACCGGGCTCGGTTCCGGCGCGGGTGGCGGAGCGCGAGGGGT 793

RESULT 13
US-09-283-471A-38
; Sequence 38, Application US/09283471A
; Patent No. 6340673
; GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; TITLE OF INVENTION: Method For Treating Tumorigenic Diseases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/283,471A
; FILING DATE: 04-APR-1999
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/861,233
; FILING DATE: 31-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/419,853
; FILING DATE: 11-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/483,533
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, James P.
; REGISTRATION NUMBER: 28,491
; REFERENCE/DOCKET NUMBER: 27373/32742A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1280 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-283-471A-38

Query Match 1.8%; Score 51.8; DB 4; Length 1280;
Best Local Similarity 51.0%; Pred. No. 0.034;
Matches 122; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
QY 6 GGGCTCGGGGTGAGCGCTAAGCGCCCTCCCGCGCGGGGCCCGCGCGCGCG 65
Db 555 GCGGGCGGGGAGGGGGCGGAAGCCCGCGACCCCGCGACCCCGCGACCCCGCGG 614
QY 66 CCGGCTGCCCGCGCGGCCATGGCGCTCCGCGCGGGCTGTGGCAGCGCTCTTGGGC 125

Db 615 ACCCCGCGACCCCGCGACCCCGCGACCCCGCGCGGCTGCGCTTCTGCCCCACGTC 674
QY 126 ATAGTCTCGCGCTTGGCTCGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAAAC 185
Db 675 CGGGTGGCGCACTGGTGGTCTGGGCTCGGCGCCCGGCTGGCGCGCGGCTCGTGG 734
QY 186 CCAGTGCCTGGTGCCCAACCGGACCTGCTTCCCACTTCTGCTGGAGCCCGAGGATGT 244
Db 735 GCCCGGAGCGGGCCGACCGGGCTCGGTTCCGGCGCGGGTGGCGGAGCGCGAGGGGT 793

RESULT 14
US-09-252-991A-7947
; Sequence 7947, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7947
; LENGTH: 699
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7947

Query Match 1.8%; Score 51.2; DB 4; Length 699;
Best Local Similarity 47.5%; Pred. No. 0.038;
Matches 152; Conservative 0; Mismatches 168; Indels 0; Gaps 0;
QY 1876 AGGCGACCTGGGACGAGGATGTGCTGCACCTGGCGGAGGAGCGCCCTTACTTACT 1935
Db 168 ACAGCTCTAGCAGCAACGAGGTGCTCGAGGCCAAGCTGCTCGGGGCAATTCGGCT 227
QY 1936 ACTGCCAGCTGGAGGCCAGTGCCTGTACGTCTTTCACGAGCAGCTGGCGCGCTTGGCC 1995
Db 228 ACAGCTGTGGTGCCCGACGACAGCTTCTCTGCCCACTACTGAAGCCGAGGTGTTC 287
QY 1996 TGGTGGAGAGGCCCTCAGCGTGGCTCGCGCCAGCGCTCAAGTGTCTTGTTCGCG 2055
Db 288 AGCCGCTGGACAGAGCAAGCTGCCGAACCTGGAAGAACCTCAACCGGCCCTGCTCAAG 347
QY 2056 CGGTGGCTGCACCTCCTCGAGTACAACTCCGGGTCTACTGCTGCATGACACCCACG 2115
Db 348 TGCTCGCGGCAAGGACCCCGCAACCGCTACGTGATGCCCTACATGTGGGGCACCAACG 407
QY 2116 ATGCACCTCAAGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGACAGCTGATCCAGGAC 2175
Db 408 GCATCGCTCAAACTCGACAAAGGTCCGCGCGGTGCTCGGCGACGATCGCGCGTGGACT 467
QY 2176 CACGGTCTGCACCTTCAAG 2195
Db 468 CTGGGACCTGGTGTCAAG 487

RESULT 15
US-09-252-991A-7809
; Sequence 7809, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

```
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7809
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7809

Query Match      1.8%; Score 51.2; DB 4; Length 1302;
Best Local Similarity 47.5%; Pred. NO. 0.046;
Matches 152; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

Qy 1876 AGGCAGCTGGGAGCAGGATCTGTGCACCTGGGCGAGGAGCGCCCTCCACCTCTACT 1935
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
368  ACGACGTCTACGACAGCAGCAGAGTGCTCGAGGCCAAGCTGTCTCGGGCATTCGGGCT 427
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1936 ACTGCCAGCTGGAGGCCAGTCCCTGTCTACGTCTTACCCGAGCAGCTGGGCGGCTTTGGCC 1995
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
428  ACGACCTGTGTGGTCCCGCAGCAGACTTCTGTGCCAACTACCTGAAGCCGAGGTGTCC 487
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1996 TGGTGGGAGAGGCCCTCAGCGTGTCTGCCGCCAAGCGCCTCAAGCTGTCTGTGTGGC 2055
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
488  AGCCGCTGGACAAGAGCAAGCTGCCGAACCTGGAAGAACTCAACCCGGCCCTGCTCAAG 547
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 2056 CGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCTGCATGACACCCACG 2115
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
548  TGCTCGCCGCGAAGGACCCCGGCAACCGTACGTGTATGCCCTACATGTGGGGCACCACG 607
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 2116 ATGCATCTCAAGGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGGACAGCTGATCCAGGAGC 2175
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
608  GCATCGCCTACAACCTCGACAAGTCCGCGGTGCTCGCGCAGCATGCCCGCTGGACT 667
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 2176 CACGGGTCTGTCACTCAAG 2195
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
668  CTTGGACCTGGTGTTCAG 687
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: October 6, 2004, 17:53:37
Job time : 205 secs

This Page Blank (uspto)

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	2881	100.0	2881	11	US-09-970-944-1		Sequence 1, Appli
2	2676.4	92.9	2752	13	US-09-918-779-1		Sequence 1, Appli
3	2676.4	92.9	2752	17	US-10-824-932-1		Sequence 1, Appli
4	2402.6	83.4	3580	17	US-10-311-623-13		Sequence 13, Appli
5	2235.8	77.6	3014	10	US-09-933-261-1		Sequence 1, Appli
6	2235.8	77.6	3014	15	US-10-256-702-1		Sequence 1, Appli
7	2189.4	76.0	2697	15	US-10-240-154-15		Sequence 15, Appli
8	1570.4	54.5	1787	15	US-09-933-261-2		Sequence 2, Appli
9	1570.4	54.5	1787	15	US-10-256-702-2		Sequence 2, Appli
10	1193.6	41.4	1321	13	US-10-296-115-365		Sequence 365, App
11	1147.8	39.8	1282	9	US-09-833-381-1806		Sequence 1806, Ap
12	889	30.9	2860	13	US-10-087-684-1		Sequence 1, Appli
13	889	30.9	2860	13	US-10-218-779-1		Sequence 1, Appli
14	887.4	30.8	2860	13	US-10-087-684-3		Sequence 3, Appli

121 TGGGCATAGTCTCGCCGCTTGGCTCCGGCGCTCGGGTCCCGAGCAGAGTGCCACCGTGG 180

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being predicted and is derived by analysis of the total score distribution.

Db	121	TGGCGATAGTCTCTCGCCGCTTGGCTTCCGGCTCGGGTGGCCGACGAGTGGCCACCGTGG	180
QY	181	CGAACCAGGTGCTGGTGGCCAAACCGGACCTGCTTCCCACTTCTCTGGTGGAGCCCGAGG	240
Db	181		
Db	181	CGAACCAGGTGCTGGTGGCCAAACCGGACCTGCTTCCCACTTCTCTGGTGGAGCCCGAGG	240
QY	241	ATGTGTACATCGTCAAGAAACAAGCCAGTCTGCTTGTGCAAGCCGTGCCGCCACGC	300
Db	241		
Db	241	ATGTGTACATCGTCAAGAAACAAGCCAGTCTGCTTGTGCAAGCCGTGCCGCCACGC	300
QY	301	AGATCTTCTTAAGTGCACCGGGAGTGGTGGCCAGTGGACACCTGATTCAGAGCGCA	360
Db	301		
Db	301	AGATCTTCTTAAGTGCACCGGGAGTGGTGGCCAGTGGACACCTGATTCAGAGCGCA	360
QY	361	GCACAGCGGGAGCAGTGGTGGACCCAGCATGGAGTCCGATTAATGTCTCAAGCGAGC	420
Db	361		
Db	361	GCACAGCGGGAGCAGTGGTGGACCCAGCATGGAGTCCGATTAATGTCTCAAGCGAGC	420
QY	421	AGGTTCGAGAAAGTGTTCGGGCTCGGAGAAATATCTGTTGCAGTGGTGGCATGGAGTCTCT	480
Db	421		
Db	421	AGGTTCGAGAAAGTGTTCGGGCTCGGAGAAATATCTGTTGCAGTGGTGGCATGGAGTCTCT	480
QY	481	CGGGCACACCAAGAGTCAGAAAGCCCTACATCGCATAGCCAGATTGGCGAAGAACTTCG	540
Db	481		
Db	481	CGGGCACACCAAGAGTCAGAAAGCCCTACATCGCATAGCCAGATTGGCGAAGAACTTCG	540
QY	541	AGCAGGAGCCGCTGGCCAAAGAGTGTCCCTGGAGCAGGGCATCTGTCCTGCCCTCCGCTC	600
Db	541		
Db	541	AGCAGGAGCCGCTGGCCAAAGAGTGTCCCTGGAGCAGGGCATCTGTCCTGCCCTCCGCTC	600
QY	601	CACCGAGGGCATCCCTCCAGCCGAGTGGTTCGCGAACGAGGACCTGGTGGACC	660
Db	601		
Db	601	CACCGAGGGCATCCCTCCAGCCGAGTGGTTCGCGAACGAGGACCTGGTGGACC	660
QY	661	CGTCCCTGGACCCCAATGTATACATCAGCGGGAGCACAGCTGGTGGTGGCAGAGGCC	720
Db	661		
Db	661	CGTCCCTGGACCCCAATGTATACATCAGCGGGAGCACAGCTGGTGGTGGCAGAGGCC	720
QY	721	GCCTTGCTGACACGGCCAACTPACACCTCGCTGGCCAAAGAACATCGTGGCACGCTCGCCGCA	780
Db	721		
Db	721	GCCTTGCTGACACGGCCAACTPACACCTCGCTGGCCAAAGAACATCGTGGCACGCTCGCCGCA	780
QY	781	GGCCCTCCGCTGCTGTACATCGTCTPACGTGAAACGGTGGTGGTGCAGCTGGACCGAGTGGT	840
Db	781		
Db	781	GGCCCTCCGCTGCTGTACATCGTCTPACGTGAAACGGTGGTGGTGCAGCTGGACCGAGTGGT	840
QY	841	CGGTCTGAGCGCCAGCTGTGGGCGCGCTGGCAGAAACGGAGCCGAGCTGCACCAACC	900
Db	841		
Db	841	CGGTCTGAGCGCCAGCTGTGGGCGCGCTGGCAGAAACGGAGCCGAGCTGCACCAACC	900
QY	901	CGGCGCTCTCAACGGGGCGCTTCTTGTGAGGGGCAGAAATGTCCATGACCGCACCGTCT	960
Db	901		
Db	901	CGGCGCTCTCAACGGGGCGCTTCTTGTGAGGGGCAGAAATGTCCATGACCGCACCGTCT	960
QY	961	CTCTCTGCTTGTCTCTGTGACCGGACCTGGAGCCCGTGGAGCAAGTGGTGGCCCTGTG	1020
Db	961		
Db	961	CTCTCTGCTTGTCTCTGTGACCGGACCTGGAGCCCGTGGAGCAAGTGGTGGCCCTGTG	1020
QY	1021	GGGTGAGTGCACCCACTGGCGGAGCGGTGAGTCTCTGACCCAGCACCCCGCAACGGAG	1080
Db	1021		
Db	1021	GGGTGAGTGCACCCACTGGCGGAGCGGTGAGTCTCTGACCCAGCACCCCGCAACGGAG	1080
QY	1081	GGGAGGAGTGCACAGGGGACTGACCTGGACACCCGCAACTGTACCACTGACCTCTGTGTAC	1140
Db	1081		
Db	1081	GGGAGGAGTGCACAGGGGACTGACCTGGACACCCGCAACTGTACCACTGACCTCTGTGTAC	1140
QY	1141	ACAGTGTCTTGCCCTTGAGGAGCTGGCCCTCTATGTGGGCCCTCATTCGCGTGGCCGTCT	1200
Db	1141		
Db	1141	ACAGTGTCTTGCCCTTGAGGAGCTGGCCCTCTATGTGGGCCCTCATTCGCGTGGCCGTCT	1200
QY	1201	GCTGTGCTGTGCTGTCTGTCTCATCTCTGTTATTGGCGGAAGAAAGAGGGCGCTGG	1260
Db	1201		
Db	1201	GCTGTGCTGTGCTGTCTGTCTCATCTCTGTTATTGGCGGAAGAAAGAGGGCGCTGG	1260

1261	ACTCAGATGGCTGACTGCTCCATTCTCACTCAGGCTTCCAGCCCGTCAAGTCAACG	1320
1261	ACTCAGATGGCTGACTGCTCCATTCTCACTCAGGCTTCCAGCCCGTCAAGTCAACG	1320
1261	ACTCAGATGGCTGACTGCTCCATTCTCACTCAGGCTTCCAGCCCGTCAAGTCAACG	1320
1321	CCAGCAAGCAGACAACCCCAATCTGCTCAACATCCAGCCGGAAGCTCAGACCAACCA	1380
1321	CCAGCAAGCAGACAACCCCAATCTGCTCAACATCCAGCCGGAAGCTCAGACCAACCA	1380
1321	CCAGCAAGCAGACAACCCCAATCTGCTCAACATCCAGCCGGAAGCTCAGACCAACCA	1380
1381	CCTACCAAGGCGAGTCTCTGTCTCCCGCGAGGATGGGCGCCAGCCCAAGTTCCAGCTC	1440
1381	CCTACCAAGGCGAGTCTCTGTCTCCCGCGAGGATGGGCGCCAGCCCAAGTTCCAGCTC	1440
1381	CCTACCAAGGCGAGTCTCTGTCTCCCGCGAGGATGGGCGCCAGCCCAAGTTCCAGCTC	1440
1441	ATGGGCACTGCTCAGCCCGCTGGGTGGCGCGCCACACACTGCACCAAGCTCTCCCA	1500
1441	ATGGGCACTGCTCAGCCCGCTGGGTGGCGCGCCACACACTGCACCAAGCTCTCCCA	1500
1441	ATGGGCACTGCTCAGCCCGCTGGGTGGCGCGCCACACACTGCACCAAGCTCTCCCA	1500
1501	CCTCTGAGGCCGAGGAGTTGCTCTCCCGCTCTCCACCCAGAACTACTTTCGCTCCCTGC	1560
1501	CCTCTGAGGCCGAGGAGTTGCTCTCCCGCTCTCCACCCAGAACTACTTTCGCTCCCTGC	1560
1501	CCTCTGAGGCCGAGGAGTTGCTCTCCCGCTCTCCACCCAGAACTACTTTCGCTCCCTGC	1560
1561	CCGAGGCCACCAACATGACCTATGGGACCTTCAACTTTCCTCGGGGCGCGGCTGATGA	1620
1561	CCGAGGCCACCAACATGACCTATGGGACCTTCAACTTTCCTCGGGGCGCGGCTGATGA	1620
1561	CCGAGGCCACCAACATGACCTATGGGACCTTCAACTTTCCTCGGGGCGCGGCTGATGA	1620
1621	TCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCT	1680
1621	TCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCT	1680
1621	TCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCT	1680
1681	ATGAGATCTACTTCAAGCTGTCACAAGCCGGAAGAGTGAGGTTGCCCTAGCTGCTGTC	1740
1681	ATGAGATCTACTTCAAGCTGTCACAAGCCGGAAGAGTGAGGTTGCCCTAGCTGCTGTC	1740
1681	ATGAGATCTACTTCAAGCTGTCACAAGCCGGAAGAGTGAGGTTGCCCTAGCTGCTGTC	1740
1741	AGACCTGCTGAGTCCCATCGTTAGCTGTGGACCCCTCGCGTCTGTCTACCCGCGCAG	1800
1741	AGACCTGCTGAGTCCCATCGTTAGCTGTGGACCCCTCGCGTCTGTCTACCCGCGCAG	1800
1741	AGACCTGCTGAGTCCCATCGTTAGCTGTGGACCCCTCGCGTCTGTCTACCCGCGCAG	1800
1801	TCATCCTGGCTATGGACAACATGTGGGAGCCAGCCCTCAGACGTGGAGCTGGCCCTCA	1860
1801	TCATCCTGGCTATGGACAACATGTGGGAGCCAGCCCTCAGACGTGGAGCTGGCCCTCA	1860
1801	TCATCCTGGCTATGGACAACATGTGGGAGCCAGCCCTCAGACGTGGAGCTGGCCCTCA	1860
1861	AAAAGCAGTGTGGGAGGCGAGCTGGGAGCAGATGTGTGACCTGGCGAGGAGGCGC	1920
1861	AAAAGCAGTGTGGGAGGCGAGCTGGGAGCAGATGTGTGACCTGGCGAGGAGGCGC	1920
1861	AAAAGCAGTGTGGGAGGCGAGCTGGGAGCAGATGTGTGACCTGGCGAGGAGGCGC	1920
1921	CCTCCACCTCTACTCTGCGCAGCTGGAGGCCAGTGCCCTGCTACGTCTTCAACGAGCAGC	1980
1921	CCTCCACCTCTACTCTGCGCAGCTGGAGGCCAGTGCCCTGCTACGTCTTCAACGAGCAGC	1980
1921	CCTCCACCTCTACTCTGCGCAGCTGGAGGCCAGTGCCCTGCTACGTCTTCAACGAGCAGC	1980
1981	TGGGCGCTTTGGCCCTGGTGGGAGAGCCCTCAGCGTGCTGGCGCCCAAGCGCTCAAGC	2040
1981	TGGGCGCTTTGGCCCTGGTGGGAGAGCCCTCAGCGTGCTGGCGCCCAAGCGCTCAAGC	2040
1981	TGGGCGCTTTGGCCCTGGTGGGAGAGCCCTCAGCGTGCTGGCGCCCAAGCGCTCAAGC	2040
2041	TGCTTCTGTTTGGCGCGTGGCCCTGCACTCCCTTCGAGTAAACATCCGGGTCTACTGCC	2100
2041	TGCTTCTGTTTGGCGCGTGGCCCTGCACTCCCTTCGAGTAAACATCCGGGTCTACTGCC	2100
2041	TGCTTCTGTTTGGCGCGTGGCCCTGCACTCCCTTCGAGTAAACATCCGGGTCTACTGCC	2100
2101	TGCATGACACCAAGATGCACTAAGAGAGTGTGACGTGGAGAGCAGCTGGGGGAGC	2160
2101	TGCATGACACCAAGATGCACTAAGAGAGTGTGACGTGGAGAGCAGCTGGGGGAGC	2160
2101	TGCATGACACCAAGATGCACTAAGAGAGTGTGACGTGGAGAGCAGCTGGGGGAGC	2160
2161	AGCTGATCCAGGAGCCAGGGTCTGCACTTCAAGGACAGTTACCAAACTCGCGCTAT	2220
2161	AGCTGATCCAGGAGCCAGGGTCTGCACTTCAAGGACAGTTACCAAACTCGCGCTAT	2220
2161	AGCTGATCCAGGAGCCAGGGTCTGCACTTCAAGGACAGTTACCAAACTCGCGCTAT	2220
2221	CCATCCAGATGTGCCAGCTCCCTGTGGAAAGATGAGCTCTTGTGAGCTAACAGAGA	2280
2221	CCATCCAGATGTGCCAGCTCCCTGTGGAAAGATGAGCTCTTGTGAGCTAACAGAGA	2280
2221	CCATCCAGATGTGCCAGCTCCCTGTGGAAAGATGAGCTCTTGTGAGCTAACAGAGA	2280
2281	TCCCCTTTTATCAATCTGGAATGCGACGAGCGGTACTTTCGACTGCACCTTCAACCTGG	2340
2281	TCCCCTTTTATCAATCTGGAATGCGACGAGCGGTACTTTCGACTGCACCTTCAACCTGG	2340
2281	TCCCCTTTTATCAATCTGGAATGCGACGAGCGGTACTTTCGACTGCACCTTCAACCTGG	2340

[illegible]

RESULT 3

US-10-624-932-1

[illegible]

QY	1362	GACCTCAG--CACCACACCACTACAGGGCAGTCTCTGTCCCGCGCAGGATGGGGCC	1418
Db	1318	GACCTCAGCACCACCAACCACTACACAGGAGTCTCTGTCCCGCGCAGGATGGGGCC	1377
QY	1419	AGCCCCAAGTTCCAGCTCACCAATGGGCACTCTGTCAAGCCCCCTGGGTGGCGGCGCCAC	1478
Db	1378	AGCCCCAAGTTCCAGCTCACCAATGGGCACTCTGTCAAGCCCCCTGGGTGGCGGCGCCAC	1437
QY	1479	ACACTGCACACAGCTCTCCACCTCTGAGGCGGAGGAGTGGTCTCCCGGCTCTCCACC	1538
Db	1438	ACACTGCACACAGCTCTCCACCTCTGAGGCGGAGGAGTGGTCTCCCGGCTCTCCACC	1497
QY	1539	CAGAACTACTTTCGGCTCCCTGCCCGGAGGACACAGCAACATGACCTATGGGACCTTCAAC	1598
Db	1498	CAGAACTACTTTCGGCTCCCTGCCCGGAGGACACAGCAACATGACCTATGGGACCTTCAAC	1557
QY	1599	TTCTCTGGGGGCGGGTGATGATCCCTAATAACAGGTATCAGGCTCTCATCCCCCAGAT	1658
Db	1558	TTCTCTGGGGGCGGGTGATGATCCCTAATAACAGGTATCAGGCTCTCATCCCCCAGAT	1617
QY	1659	GCCATACCCCGAGGGAAGATCTATGAGATCTACTCAGCTGCACAAGCGGAAGACGTG	1718
Db	1618	GCCATACCCCGAGGGAAGATCTATGAGATCTACTCAGCTGCACAAGCGGAAGACGTG	1677
QY	1719	AGGTTGCCCTAGCTGSGCTCTCAGACCGCTGCTGAGTCCCATCGTTAGCTGTGGACCCCT	1778
Db	1678	AGGTTGCCCTAGCTGSGCTCTCAGACCGCTGCTGAGTCCCATCGTTAGCTGTGGACCCCT	1737
QY	1779	GGCGTCTCTCACCOCGCCAGTCACTCTGGCTATGGACCACTGTGGGAGGCCACAGCCCT	1838
Db	1738	GGCGTCTCTCACCOCGCCAGTCACTCTGGCTATGGACCACTGTGGGAGGCCACAGCCCT	1797
QY	1839	GACAGCTGGAGCCCTGCCCTCAAAAAGCAGTCTGTGGAGGCGAGCTGGGAGCAGGATGTG	1898
Db	1798	GACAGCTGGAGCCCTGCCCTCAAAAAGCAGTCTGTGGAGGCGAGCTGGGAGCAGGATGTG	1854
QY	1899	CTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACTATCTGACAGCTGGAGGCCAGTGCC	1958
Db	1855	CTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACTATCTGACAGCTGGAGGCCAGTGCC	1914
QY	1959	TGCTACGCTCTTACCGAGCAGCTGGGCGGCTTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	2018
Db	1915	TGCTACGCTCTTACCGAGCAGCTGGGCGGCTTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG	1974
QY	2019	GCTGCCGCAAGCGCCCTAAGCTGCTTCTGTTTGGCCCGGTGGCGCTGCACCTCCCTCGAG	2078
Db	1975	GCTGCCGCAAGCGCCCTAAGCTGCTTCTGTTTGGCCCGGTGGCGCTGCACCTCCCTCGAG	2034
QY	2079	TACAACATCCGGGTCTACTGCTCGATGACACCAACATGCTACTCAGAGGAGTGGTGCAG	2138
Db	2035	TACAACATCCGGGTCTACTGCTCGATGACACCAACATGCTACTCAGAGGAGTGGTGCAG	2094
QY	2139	CTGGAGAACAGCTGGGGGGAAGCTGATCCAGAGGCCACGGGTCTGTCACTTCAAGGAC	2198
Db	2095	CTGGAGAACAGCTGGGGGGAAGCTGATCCAGAGGCCACGGGTCTGTCACTTCAAGGAC	2154
QY	2199	AGTTACCAACAACCTGGCGCTATCCATCCAGATGTGCCAGCTCCCTGTGGAGAGATTAAG	2258
Db	2155	AGTTACCAACAACCTGGCGCTATCCATCCAGATGTGCCAGCTCCCTGTGGAGAGATTAAG	2214
QY	2259	CTCCCTTGTGAGTACAGGAGATCCCTTTTATCACAATCTGGAATGCACGCGACGGTAC	2318
Db	2215	CTCCCTTGTGAGTACAGGAGATCCCTTTTATCACAATCTGGAATGCACGCGACGGTAC	2274
QY	2319	TTCGACTGCACTTCAACCTGGAGCGTGTACGCCCCAGACCTAGTGACCTGGCTGCAAG	2378
Db	2275	TTCGACTGCACTTCAACCTGGAGCGTGTACGCCCCAGACCTAGTGACCTGGCTGCAAG	2334
QY	2379	CTGTGGGTGTGCAGGTGGAGGGCCACGGGAGAGCTTTCAGCATCAACTTCAACATCAAC	2438
Db	2335	CTGTGGGTGTGCAGGTGGAGGGCCACGGGAGAGCTTTCAGCATCAACTTCAACATCAAC	2394

RESULT 4

RESULTS 4
IIS-10-311-623-13US-10-311-623-13
Application US/10311623

```

; Sequence 13, Application US/103

```

; Publication No. US20

; GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.

APPLICANT: KALLICK, Deborah A.; TRIBOULEY

APPLICANT: YUE, Henry; NGUYEN, Danniell B.

APPLICANT: TANG, Y. Tom: LAL. Preeti G

APPLICANT: POLICKY JENNIFER L. : AZIMZAT VALDA

APPLICANT: POLICKY, JENNIFER L.; AZIMZAI, FARUQ
 AZIMZAI, FARUQ M.; CRAN, RICHARD C

APPLICANT: LU, Dyung Alna M.; GRAUL, Richard C.

; APPLICANT: YAO, Monique G.; BURFORD, Neil

APPLICANT: HAFALIA, April J. A.; BAUGHN, Mariah

APPLICANT: BANDMAN, Olga; ARVIZU, Chandra S.

APPLICANT: YANG, Junming; XU, Yuming

APPLICANT: GANDHI, Ameena R. : WARREN, Bridget

APPLICANT: DING I-I: SANTIANWAI-A. Madhusudan M. GANDHI, AMEETHA K.; NARAYAN, EUGENE

APPLICANT: DING, LI; SANJANWALA, Maheshwari M.
 ADDRESS: DUCAN Road M. J. Van

APPLICANT: DUGGAN, Brendan M.; LU, Yail

; TITLE OF INVENTION: RECEPTORS

; FILE REFERENCE: PF-0793 USN

; CURRENT APPLICATION NUMBER: US/10/311,623

CURRENT FILING DATE: 2002-12-17

PRIOR APPLICATION NUMBER: US 01/19942

; PRIOR APPLICATION NUMBER: 03 01/15512
 ; PRIOR FILING DATE: 2001-06-21

; PRIOR FILING DATE: 2001-06-21
 ; PRIORITY CITATION NUMBER: US 60/214 027

; PRIOR APPLICATION NUMBER: US 60/214,027
 ; FILING DATE: 2000-06-03

; PRIOR FILING DATE: 2000-06-21

; PRIOR APPLICATION NUMBER: US 60/228,045

;
PRIOR FILING DATE: 2000-08-25

PRIOR APPLICATION NUMBER: US 60/255,104

PRIOR FILING DATE: 2000-12-12

; PRIOR FILING DATE: 2000-12-12
 ; NUMBER OF SEQ ID NOS: 24

```
; NUMBER OF SEQ ID NOS: 24
```

```

; SOFTWARE: PERL Program
;

```

; SEQ ID NO 13

LENGTH: 3580

TYPE: DNA

Case	Age	Score	DB	Length
1	17	3580		
2	17	3580		
3	17	3580		
4	17	3580		
5	17	3580		
6	17	3580		
7	17	3580		
8	17	3580		
9	17	3580		
10	17	3580		
11	17	3580		
12	17	3580		
13	17	3580		
14	17	3580		
15	17	3580		
16	17	3580		
17	17	3580		
18	17	3580		
19	17	3580		
20	17	3580		
21	17	3580		
22	17	3580		
23	17	3580		
24	17	3580		
25	17	3580		
26	17	3580		
27	17	3580		
28	17	3580		
29	17	3580		
30	17	3580		
31	17	3580		
32	17	3580		
33	17	3580		
34	17	3580		
35	17	3580		
36	17	3580		
37	17	3580		
38	17	3580		
39	17	3580		
40	17	3580		
41	17	3580		
42	17	3580		
43	17	3580		
44	17	3580		
45	17	3580		
46	17	3580		
47	17	3580		
48	17	3580		
49	17	3580		
50	17	3580		
51	17	3580		
52	17	3580		
53	17	3580		
54	17	3580		
55	17	3580		
56	17	3580		
57	17	3580		
58	17	3580		
59	17	3580		
60	17	3580		
61	17	3580		
62	17	3580		
63	17	3580		
64	17	3580		
65	17	3580		
66	17	3580		
67	17	3580		
68	17	3580		
69	17	3580		
70	17	3580		
71	17	3580		
72	17	3580		
73	17	3580		
74	17	3580		
75	17	3580		
76	17	3580		
77	17	3580		
78	17	3580		
79	17	3580		
80	17	3580		
81	17	3580		
82	17	3580		
83	17	3580		
84	17	3580		
85	17	3580		
86	17	3580		
87	17	3580		
88	17	3580		
89	17	3580		
90	17	3580		
91	17	3580		
92	17	3580		
93	17	3580		
94	17	3580		

Query Match 83.4%; Score 2402.6; DB 17; Length 3360;

Best Local Similarity	93.4%	Pred. No. 0;	177.
-----------------------	-------	--------------	------

[illegible]

1 GCGATGGCCGTCGGCCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCTCGCCGCTTG 60

[illegible]

RESULT 5

US-09-933-261-1
Sequence 1, Application US/09933261
Publication No. US20030040046A1
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
Leonardo, E. David
Hink, Lindsay
Masu, Masayuki
Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/933,261
FILING DATE: 20-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/808,982
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A

NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 STREET: 268 BUSH STREET, SUITE 3200
 CITY: SAN FRANCISCO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/256,702
 FILING DATE: 27-Sep-2002
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/933,261
 FILING DATE: 20-Aug-2001
 APPLICATION NUMBER: 08/808,982
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: OSMAN, RICHARD A.
 REGISTRATION NUMBER: 36,627
 REFERENCE/DOCKET NUMBER: UC96-217
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 343-4341
 TELEFAX: (415) 343-4342
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3014 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-10-256-702-1

Query Match 77.6%; Score 2235.8; DB 15; Length 3014;
 Best Local Similarity 88.5%; Pred. No. 0;
 Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;

Qy	87	ATGCGCGTCCGGCCCGGCTGTGGCAGTCTCTGGGCATAGTCTCTCGCGCGCTGGCTT	146
Db	1	ATGCGCGTCCGGCCCGGCTGTGGCAGTCTCTGGGCATAGTCTCTCGCGCGCTGGCTT	60
Qy	147	CGCGCGTCCGGTCCGACGAGTGCCACCGTGGCCAAACCCAGTCCCTGGTGCCAAACCCG	206
Db	61	CGTGGTCCGGTCCGACGAGTGCCACCGTGGCCAAACCCAGTCCCTGGTGCCAAACCCG	120
Qy	207	GACCTGTCTCCCACTTCTCTGGTGAGCCCGAGGATGTATCGTCAAGAAACAAAGCCA	266
Db	121	GACCTGTCTCCCACTTCTCTGGTGAGCCCTGAGGACGTGTATGTCAAGAAACAAAGCCG	180
Qy	267	GTGCTGTCTGTGCAAGCCGTCGCCCGCAGCAGATCTTCTCAAGTGCAACCGGGAG	326
Db	181	GTGTTGTGTGTGCAAGCCGTCGCCCGCAGCAGATCTTCTCAAGTGCAATGGGAA	240
Qy	327	TGGGTGCCAGGTGGACCACTGATCGAGCGCAGCAGCGGAGCAGTGGTGAGCCG	386
Db	241	TGGGTGCCAGGTGGACCACTGATCGAGCGCAGCAGCGGAGCAGTGGTGAGCCG	300
Qy	387	ACCATGAGGTCCCGATTAATGTCTCAAGGAGCAGGTGCGAAGGTGTTCGGGCTGGAG	446
Db	301	ACCATGAGGTCCCGATTAATGTCTCAAGGAGCAGGTGCGAAGGTGTTCGGGCTGGAG	360
Qy	447	GAATAGTGTGCGAGTGGCATGGAGTCTCTCGGCGACCAACCAAGAGTCAGAGGCC	506
Db	361	GAATAGTGTGCGAGTGGCATGGAGTCTCTCGGCGACCAACCAAGAGTCAGAGGCC	420
Qy	507	TACATCCGATGACCAAGATTTGGCGAAGAACTTCGAGCAGGACCCGCTGGCCAAAGAGGTG	566

Db	421	TACATCCGATTTGCTTAITTTGGCAAGAACTTTTGACAGAGCCACTGGCCAAAGGAAGTG	480
Qy	567	TCCCTGGACAGGGGATCTGTGCTGCCCTGCCGTCACACCGAGGGCATCTCCCTCCAGCCGAG	626
Db	481	TCACTGGACAAAGGCAITGTACTCTTGTGCCCGCCAGAGGAATCCCCCAGCTGAG	540
Qy	627	GTGGAGTGGCTCCGGAACGAGGACCTGTGTGGACCCCTCCCTGGACCCCAATGTATCATC	686
Db	541	GTGGAGTGGCTTCGAATGAGGACCTGTGTGGACCCCTCCCTCGATCCCAATGTATCATC	600
Qy	687	ACCGGGAGCAGACCTGTGTGTGGACAGGCGCGCTTGTCTGACAGGCGCAACTACAC	746
Db	601	ACCGGGAGCAGACCTGTGTGTGGACAGGCGCGCTTGTCTGACAGGCGCAACTACAC	660
Qy	747	TGCGTGGCCAAAGAACATCGTGGCACCTGCGCGAGCGCTCCCGTGTGTCTCATCTAC	806
Db	661	TGTGTGGCCAAAGAACATCGTAGCCCGTGTGGCGAAGCACTCTGAGCGGTCATGTTAT	720
Qy	807	GTGAACGGTGGTGGTGGACCGAGTGTGTGGTGTGGACCGCAGCTGTGGCGGC	866
Db	721	GTGAACGGTGGTGGTGGACCGAGTGTGTGGTGTGGACCGCAGCTGTGGCGGT	780
Qy	867	GGCTGGCAGAAACGAGCGCGAGCTGCACCAACCGCGGCTCTCAACGGGGCGCTTC	926
Db	781	GGCTGGCAGAAACGAGCGCGAGCTGCACCAACCGCGGCTCTCAACGGGGCGCTTC	840
Qy	927	TGTGAGGGCAGAAATGTCATGACCGCACCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	986
Db	841	TGTGAGGGCAGAAATGTCAGAAAAACAGCTGTGGCGCACTCTG---TGCCAGTGGATGG	897
Qy	987	AGCTTGAGCCCGTGGAGCAAGTGTGTGGCTGTGGGTGTGGACGTGCAACCACTGGCGAGC	1046
Db	898	AGCTTGAGTTCGTGGAGTAAGTGTGTGGCTGTGGGTGTGGACGTGCAACCACTGGCGAGC	957
Qy	1047	CGT	1106
Db	958	CGCGAGTGTCTGTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTGGGGTGTGACCTG	1017
Qy	1107	GACACCGCAACTGTACAGTGACCTCTGTGTACACAGTGTCTTGTGGCCCTGAGGACGTG	1166
Db	1018	GACACCGCAACTGTACAGTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1077
Qy	1167	GCCCTCTATGTGGGCTCATCGCGTGTGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1226
Db	1078	GCTCTCTACATCGGCTTGT	1137
Qy	1227	ATCTCGTGTATTCGCGAAGAGAGGGGCTGACCTCAGATGTGGTGTGTGTGTGTGTGTGT	1286
Db	1138	GGACTCATTTACTGTGCAAGAGAGGGGCTGACCTCCGATGTGGCGGATGTGTGTGTGT	1197
Qy	1287	CTCACCTCAGGCTTCCAGCCCGTCAAGATCAAGCCCAAGCAAGAGAGAGAGAGAGAGAG	1346
Db	1198	CTCACCTCAGGCTTCCAGCCCGTCAAGATCAAGCCCAAGCAAGAGAGAGAGAGAGAG	1257
Qy	1347	CTCACCTCAGGCTTCCAGCCCGTCAAGATCAAGCCCAAGCAAGAGAGAGAGAGAGAG	1403
Db	1258	CTCACCTCAGGCTTCCAGCCCGTCAAGATCAAGCCCAAGCAAGAGAGAGAGAGAGAG	1317
Qy	1404	CGGAGGATGGCGCCAGCCCAAGTTCAGCTCAGATGTGGGCTGTGTGTGTGTGTGTGTGT	1463
Db	1318	AGGAGGATGGCGCCAGCCCAAGTTCAGCTCAGATGTGGGCTGTGTGTGTGTGTGTGTGT	1377
Qy	1464	GGTGGCGCCGCCACACACTGCAACACAGCTCTCCACCTCTGTGGCGCGAGAGTGTGTGT	1523
Db	1378	GGGAGTGGCGCCATACCTTGCACACAGCTCACCCACTCTGTGGGCTGAGGACTGTGTGT	1437
Qy	1524	TCCCGCTCTCCACCGAATATCTTCGCTTCCCTGCGCGGAGGACACAGCAACATGAC	1583
Db	1438	TCCCGCTCTCCACCGAATATCTTCGCTTCCCTGCGCGGAGGACACAGCAACATGAC	1497
Qy	1584	TATGGAGCTTCACTCTCGGGGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1643
Db	1498	TACGGGACCTTCACTTCTCTCGGGGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT	1557

2010 CTACGCTGGCTGCGCCAAAGCGCTCAAGCTGCTTCTGTTTGGCCGCGTGCCTGCACC 2069
|
|
|
895 CTCAGCTGGCTGCGCCGCAAGCGCTCAAGCTGCTTCTGTTTGGCCGCGTGCCTGCACC 954
|
|
|
2070 TCCCTCGAGTCAACATCCGGGTCTACTGCTGCAATGACACCCAGCATGCACTCAAGGAG 2129
|
|
|
955 TCCCTCGAGTCAACATCCGGGTCTACTGCTGCAATGACACCCAGCATGCACTCAAGGAG 1014
|
|
|
2130 GTGTGTCAGCTGGAGAACGCTGGGGGACAGCTGATCCAGGAGCCAGCGGTCTGTCAC 2189
|
|
|
1015 GTGTGTCAGCTGGAGAACGCTGGGGGACAGCTGATCCAGGAGCCAGCGGTCTGTCAC 1074
|
|
|
2190 TTCAGGACAGTTTACCAACCTTGGCTATCCATCCACGATGTCGCCAGCTCCCTGFGG 2249
|
|
|
1075 TT-AAGGACAGTTTACCAACCT--GCCCTATCATCCAGATGTCGCCAGCTCCCTGFGG 1131
|
|
|
2250 AAGAGTAAGCTTCTGTGTCAGTACAGGAGATCCCTTTTATCATCTGGAATGGCAGC 2309
|
|
|
1132 AAGAGTAAGCTTCTGTGTCAGTACAGGAGATCCCTTTTATCATCTGGAATGGCAGC 1191
|
|
|
2310 CAGCGGTACTTGCACTGACCTTCCACCTGGAGCGTGTGACCCCGCAGCACTAGTGACCTG 2369
|
|
|
1192 CAGCGGTACTTGCACTGACCTTCCACCTGGAGCGTGTGACCCCGCAGCACTAGTGACCTG 1251
|
|
|
2370 GCTGCAAGCTGTGGGTGTGTCAGGTGGAGGCGACGCGGACAGCTTTCAGCATCAACTTC 2429
|
|
|
1252 GCCTGCAAGCTGTGGGTGTGTCAGGTGGAGGCGACGCGGACAGCTTTCAGCATCAACTTC 1311
|
|
|
2430 AACATCACCAGGACACAAAGTTTGTGAGTGTGCTGTGAGAGTGGAAGCGGGGTC 2489
|
|
|
1312 AACATCACCAGGACACAAAGTTTGTGAGTGTGCTGTGAGAGTGGAAGCGGGGTC 1371
|
|
|
2490 CCAGCCCTGTGGGCGCCAGTGTGCTTCAAGATCCCTTCTCATTCGGCAGAGATAATT 2549
|
|
|
1372 CCAGCCCTGTGGGCGCCAGTGTGCTTCAAGATCCCTTCTCATTCGGCAGAGATAATT 1431
|
|
|
2550 TCCAGCTGTGACCCACCTGTAGGCGGGGTGCGGACTTGGCGGACTTGGGCCAGAACTC 2609
|
|
|
1432 TCCAGCTGTGACCCACCTGTAGGCGGGGTGCGGACTTGGCGGACTTGGGCCAGAACTC 1491
|
|
|
2610 CACTGTGAGCCACTCTAGCTTCTTTGCTTCCAAAGCCAGCCCGCAGCCATGATCCTC 2669
|
|
|
1492 CACTGTGAGCCACTCTAGCTTCTTTGCTTCCAAAGCCAGCCCGCAGCCATGATCCTC 1551
|
|
|
2670 AACCTGTGGAGCGCGGCACTTCCCAAGCGCAACTCAGCCAGCTGGCTGAGCAGTG 2729
|
|
|
1552 AACCTGTGGAGCGCGGCACTTCCCAAGCGCAACTCAGCCAGCTGGCTGAGCAGTG 1611
|
|
|
2730 GTTGGACTGGCCAGCCAGACGCTGGCCTC-TTCACAGTG-TGGAGGCTGAGTGTGAG 2787
|
|
|
1612 GTTGGACTGGCCAGCCAGACGCTGGCCTC-TTCACAGTG-TGGAGGCTGAGTGTGAG 1671
|
|
|
2788 GCGGCGAGCCCGACACCTACTCTACAGCTTTGGCA--CCACCAAGGACAGGCA 2845
|
|
|
1672 GCGGCGAG--CGAACACTCAAAATTTTACAGTTTGGGAAACCCACCAAGGACAGGCA 1729
|
|
|
2846 GAAGCCGACAGCGGCGCTTCCCAACACCGGGGAGA 2881
|
|
|
1730 GAAGCCGACAGCGGCTTTTCCAAACCGGGGAGA 1765
|
|
|

RESULT 10

US-10-296-115-365
; Sequence 365, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725

; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478

; SEQ ID NO 365
; LENGTH: 1321
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-296-115-365

Query Match 41.4%; Score 1193.6; DB 13; Length 1321;
Best Local Similarity 97.8%; Pred. No. 2.6e-287;
Matches 1295; Conservative 0; Mismatches 19; Indels 10; Gaps 8;
QY 1476 CACACACTGACACACAGCTCTCCACCTCTGAGGCGGAGGAGTTGCTCTCCGCGCTCTCC 1535
|
|
|
Db 1 CACACACTGACACACAGCTCTCCACCTCTGAGGCGGAGGAGTTGCTCTCCGCGCTCTCC 60
|
|
|
QY 1536 ACCAGAACTACTTCCGCTCTCCGCGGAGGACACAGCAACATGACTATGGAGCCTTC 1595
|
|
|
Db 61 ACCAGAACTACTTCCGCTCTCCGCGGAGGACACAGCAACATGACTATGGAGCCTTC 120
|
|
|
QY 1596 AACTTCTCCGCGGCGCGCTGATGATCCCTAAATACAGGTATCAGCCTCTCTATCCGCCA 1655
|
|
|
Db 121 AACTTCTCCGCGGCGCGCTGATGATCCCTAAATACAGGAATCAGCCTCTCTATCCGCCA 180
|
|
|
QY 1656 GATGCCATACCCCGAGGGAAGATCTATGAGATCTACTCAGCTGCAACAAGCCGGAAGAC 1715
|
|
|
Db 181 GATGCCATACCCCGAGGGAAGATCTATGAGATCTACTCAGCTGCAACAAGCCGGAAGAC 240
|
|
|
QY 1716 GTGAGTTGCCCTTAGCTGTGCTGCTGAGCCCTGAGTCCCATGCTTAGCTGTGGACCC 1775
|
|
|
Db 241 GTGAGTTGCCCTTAGCTGTGCTGCTGAGCCCTGAGTCCCATGCTTAGCTGTGGACCC 300
|
|
|
QY 1776 CCT-GGCGTCTCTCTCACCAGCGGCGAGTCACTCT-GGCTATGGACCACTGT-GGGAGCCC 1832
|
|
|
Db 301 CCTGGCGTCTCTCTTACCCGCGCAGTCACTCTGGGGTATGGACCACTGTGGGGAGCCCC 360
|
|
|
QY 1833 AGCCCTGACAGCT-GGAGCCTGCGCTCAAAAAGCAGTCTGCGAGGCGAGCTGGGAGCA 1891
|
|
|
Db 361 AGCCCTGACAGCTGGGAGCCTGCGCTCAAAAAGCAGTCTGCGAGGCGAGCTGGG--A 417
|
|
|
QY 1892 GATGTGCTGCACTGGGCGAGGAGCGCCCTCCACCTCTACTACTGCGAGCTGGAGGC 1951
|
|
|
Db 418 GATGTGCTGCACTGGGCGAGGAGCGCCCTCCACCTCTACTACTGCGAGCTGGAGGC 477
|
|
|
QY 1952 CAGTGTCTGTAGCTTTCACCGAGCAGCTGGGCGCTTTGGCCCTGCTGGTGGGAGGCCCT 2011
|
|
|
Db 478 CAGTGTCTGTAGCTTTCACCGAGCAGCTGAGCGCCTATGCGCTTGGTGGGAGGCCCT 537
|
|
|
QY 2012 CAGCGTGTGCGCCCAAGCGCTCAAGCTGCTTCTGTTTGGCGCGTGGCCTGCACTTC 2071
|
|
|
Db 538 CAGCGTGTGCGCCCAAGCGCTCAAGCTGCTTCTGTTTGGCGCGTGGCCTGCACTTC 597
|
|
|
QY 2072 CTTGAGTACACATCCGGGTCTACTGCTGATGACACCCAGATGCACTCAAGAGGT 2131
|
|
|
Db 598 CTTGAGTACACATCTACTGCTGATGACACTCAGATGCACTCAAGTCACTCAAGTACT 657
|
|
|
QY 2132 GGTGAGCTGGAGAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTCTCGACTT 2191
|
|
|
Db 658 GGTGAGCTGGAGAGCAGCTGCGGGACAGCTGATCCAGGAGCCACGGTCTCGACTT 717
|
|
|
QY 2192 CAAGGACAGTTACCAACCTTCGCGCTTATCCATCCACGATGTGCCAGCTCCCTCTGGAA 2251
|
|
|
Db 718 CAAGGACAGTTACCAACCTTCGCGCTTATCCATCCACGATGTGCCAGCTCCCTCTGGAA 777
|
|
|
QY 2252 GAGTAAGCTCTTGTGACGTACAGGAGATCCCTTTTATCATCTGGAATGGACGCA 2311
|
|
|
Db 778 GAGTAAGCTCTTGTGACGTACAGGAGATCCCTTTTATCATCTGGAATGGACGCA 837
|
|
|
QY 2312 GCGGTACTTGCACTGCACTTTCACCTCGAGGCTGTGACCCCGACACTAGTACCTGGC 2371
|
|
|
Db 838 GCGGTACTTGCACTGCACTTTCACCTCGAGGCTGTGACCCCGACACTAGTACCTGGC 897

QY	2372	CTGCAAGCTGTGGGTGGCAGAGTGGAGGGCGACGGGCGAGAGCTTCAGCATCAACTTCAA	2431	QY	1690	ACCTCAGCTGCACAGCCGGAAGCGTGTGAGGTTGCCCTAGCTGGCTGTCTCAGACCCCTGC	1749
Db	898	CTGCAAGCTGTGGGTGGCAGAGTGGAGGGCGACGGGCGAGAGCTTCAGCATCAACTTCAA	957	Db	181	ACCTCAGCTGCACAGCCGGAAGCGTGTGAGGTTGCCCTAGCTGGCTGTCTCAGACCCCTGC	240
QY	2432	CATCACCAAGGACACAAAGTTTCTGAGCTGTGGCTGTGGAGAGTGAAGCGGGGTGCC	2491	QY	1750	TGAGTCCCATCGTTAGCTGTGGACCCCTT-GGCGTCTCTGTCTACCCGGGCGAGTCACTCT-	1807
Db	958	CATCACCAAGGACACAAAGTTTCTGAGCTGTGGCTGTGGAGAGTGAAGCGGGGTGCC	1017	Db	241	TGAGTCCCATCGTTAGCTGTGGACCCCTTGTCTACCCGGGCGAGTCACTCTG	300
QY	2492	AGCCCTGTGGGCCCCCAGTGCCTTCAAGATCCCTTCTCTCATTCGGGAGAGATAATTTTC	2551	QY	1808	GGCTATGGAACCACTGTGGGAGCCCAAGCCCTGACAGCT-GGAGCTTGGCGCTCAAAAGC	1866
Db	1018	AGCCCTGTGGGCCCCCAGTGCCTTCAAGATCCCTTCTCTCATTCGGGAGAGATAATTTTC	1077	Db	301	GGGTATGGAACCACTGTGGGAGCCCAAGCCCTGACAGCTGGGAGCCTGCGCTCAAAAGC	360
QY	2552	CAGCCTGGACCCACCTGTAGGGGGGTGCGGACTGCGGACTCTGGCCCGAGAACTCCA	2611	QY	1867	AGTCGTGGGAGGCGAGCTGGGAGCAGATGTGCT--TGACCTTGGGCGAGAGCGCCCTCCC	1926
Db	1078	CAGCCTGGACCCACCTGTAGGGGGGTGCGGACTGCGGACTCTGGCCCGAGAACTCCA	1137	Db	361	AGTCGTGGGAGGCGAGCTGGGAGGATGTGCT--TGACCTTGGGCGAGAGCGCCCTCCC	418
QY	2612	CTTGGACAGCATCTCAGCTTCTTTGCTCCCAAGCCCGAGCCCGAGCCATGATCCTCAA	2671	QY	1927	ACCTCTACTACTGCCAGCTGGAGGCCAGTCTGTCTAAGTCTTCAACGAGAGTGGGCC	1986
Db	1138	CCTGGACAGCATCTCAGCTTCTTTGCTCCCAAGCCCGAGCCCGAGCCATGATCCTCAA	1197	Db	419	ACCTCTACTACTGCCAGCTGGAGGCCAGTCTGTCTAAGTCTTCAACGAGAGTGGGCC	478
QY	2672	CCTGTGGAGGCGGCGACCTTCCCAACGGCAACTCAGCCAGCTGGCTCAGCAGTGGC	2731	QY	1987	GCCTTGGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCGCCAAAGCGCTCAAGCTGCTC	2046
Db	1198	CCTGTGGAGGCGGCGACCTTCCCAACGGCAACTCAGCCAGCTGGCTCAGCAGTGGC	1257	Db	479	GCTATGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCGCCAAAGCGCTCAAGCTGCTC	538
QY	2732	T-GGACTGGGCCAGCAGAGAGCTGGCTC-TTCACAGTG-TGAGAGCTGAGTCTGAGG	2788	QY	2047	TGTTTGGCGCGGTGGCTGCACCTCCCTCGAGTACAACATCCCGGTCTACTGCTGCGATG	2106
Db	1258	TGGAGCTGGGCCAGCAGAGAGCTGGCTTCTTTTCAAGTGTTCGGAGGCTGAGTCTGAGG	1317	Db	539	TGTTTGGCGCGGTGGCTGCACCTCCCTCGAGTACAACATCTGCTACTGCTGCGATG	598
QY	2789	CCGG 2792		QY	2107	ACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAGAGAGCTGGGGGAGCAGCTGA	2166
Db	1318	CCGG 1321		Db	599	ACACTCAGATGCACTCAACGTAGTGTGCAGCTGGAGAGAGCAGCTGGGGGAGCAGCTGA	658
RESULT 11				QY	2167	TCAGGAGCCACGGCTCTGCACCTTCAAGGACAGTGTACCAACCTCGCGCTATCCATCC	2226
US-09-833-381-1806				Db	659	TCCAGGAGCCACGGGTCTCTGCACCTTCAAGGACAGTGTACCAACCTCGCGCTATCCATCC	718
; Sequence 1806, Application US/09833381				QY	2227	ACGATGTGCCAGCTCCCTGTGGAAGAGTAAAGTCTTGTTCAGCTACACAGAGATCCCT	2286
; Patent No. US20020132090A1				Db	719	ACGATGTGCCAGCTCCCTGTGGAAGAGTAAAGTCTTGTTCAGCTACACAGAGATCCCT	778
; GENERAL INFORMATION:				QY	2287	TTTATCATCTGGAATGGCACGCGGCTACTTGCATGCACTTCAACCTTCAACCTTCAACCT	2346
; APPLICANT: Robison, Keith E.				Db	779	TTTATCATCTGGAATGGCACGCGGCTACTTGCATGCACTTCAACCTTCAACCTTCAACCT	838
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs				QY	2347	TCAGCCCGCAGCTAGTGCACCTGCGCTGCAAGCTGTGGCTGTGGCAGTGGAGGCGAGC	2406
; FILE REFERENCE: 5800-119				Db	839	TCAGCCCGCAGCTAGTGCACCTGCGCTGCAAGCTGTGGCTGTGGCAGTGGAGGCGAGC	898
; CURRENT APPLICATION NUMBER: US/09/833,381				QY	2407	GGCAGAGCTTCAGCATCAACTTCAACATCAACAGGACACAAGGTTTGTCTGAGCTGTGG	2466
; PRIOR FILING DATE: 2001-04-11				Db	899	GGCAGAGCTTCAGCATCAACTTCAACATCAACAGGACACAAGGTTTGTCTGAGCTGTGG	958
; PRIOR APPLICATION NUMBER: 09/516,448				QY	2467	CTCTGGAGAGTGAAGCGGGGTCCAGCCCTGTGGGCCCCCAGTGCCTTCAAGATCCCT	2526
; PRIOR FILING DATE: 2000-02-29				Db	959	CTCTGGAGAGTGAAGCGGGGTCCAGCCCTGTGGGCCCCCAGTGCCTTCAAGATCCCT	1018
; NUMBER OF SEQ ID NOS: 2050				QY	2527	TCCTCATTCGGCAGAGATGA- TTTCCAGCTTGGACCCACCTGTAGCGGGGTCCGAC	2585
; SOFTWARE: FastSeq for Windows Version 3.0				Db	1019	TCCTCATTCGGCAGAGATGAATTTTTCAGCTTGGACCCACCTGTAGCGGGGTCCGAC	1078
; SEQ ID NO 1806				QY	2586	TGGCGG- ACTCTGGCCCCAGAACTCCACTTGACAGCCATCTCAGCTTCTTTGCTCCAA	2644
; LENGTH: 1282				Db	1079	TGGCGGAACTCTGGCCCCAGAACTCCACTTGACAGCCATCTCAGCTTCTTTGCTCCAA	1138
; TYPE: DNA				QY	2645	GCCAGCCCCACAGCCATGATCTCAACTGTGGAGGCGGCGACTTCCCAAGGGCAA	2704
; ORGANISM: Homo sapiens				Db	1139	GCCAGCCCCACAGCCATGATCTCAACTGTGGAGGCGGCGACTTCCCAAGGGCAA	1198
; FEATURE:				QY	2705	CCTCAGCCAGCTGGCTGAGCAGTGGCTGAGCTGGGCGGCGAGCAGCTGCTTTCAC	2764
; NAME/KEY: misc_feature				Db	1199	CCTCAGCCAGCTGG- TGCAGCAGTGGCTGAGCTGGGC- -AGCAGAGCTGG-CTCTTCAC	1254
; LOCATION: (1)...(1282)				Query Match			
; OTHER INFORMATION: n = A,T,C or G				Best Local Similarity			
US-09-833-381-1806				Matches 1259; Conservative			
				97.7%; Pred. No. 6.7e-276;			
				0; Mismatches 18; Indels 11; Gaps 9;			
QY	1510	CGAGAGTGTGTCTCCGCTCTCCAGCAGAACTTCCAGTGGGCGGCTGATGATCCCTTAATA	1629	QY	1510	CGAGAGTGTGTCTCCGCTCTCCAGCAGAACTTCCAGTGGGCGGCTGATGATCCCTTAATA	1629
Db	1	CGAGAGTGTGTCTCCGCTCTCCAGCAGAACTTCCAGTGGGCGGCTGATGATCCCTTAATA	120	Db	61	CGAGAGTGTGTCTCCGCTCTCCAGCAGAACTTCCAGTGGGCGGCTGATGATCCCTTAATA	120
QY	1630	CAGGTATCAGCTCTCTATCCCGGCGGAGTGCATATCCCGGAGGAGATCTATGATCT	1689	QY	1630	CAGGTATCAGCTCTCTATCCCGGCGGAGTGCATATCCCGGAGGAGATCTATGATCT	1689
Db	121	CAGGTATCAGCTCTCTATCCCGGCGGAGTGCATATCCCGGAGGAGATCTATGATCT	180	Db	121	CAGGTATCAGCTCTCTATCCCGGCGGAGTGCATATCCCGGAGGAGATCTATGATCT	180

QY 2765 AGTGTGGAGGCTGAGTGTGAGGCGG 2792
 Db 1255 AGTGTGGAGGCTGAGTGTGAGGCGG 1282

RESULT 12

US-10-087-684-1
 ; Sequence 1, Application US/10087684
 ; Publication No. US20040029116A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Edinger, Shlomit R.
 ; APPLICANT: MacDougall, John R.
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Stone, David J.
 ; APPLICANT: Grosse, William M.
 ; APPLICANT: Leplev, Denise M.
 ; APPLICANT: Rieger, Daniel K.
 ; APPLICANT: Burgess, Catherine E.
 ; APPLICANT: Casman, Stacie, J.
 ; APPLICANT: Spytek, Kimberly A.
 ; APPLICANT: Boldog, Ferenc L.
 ; APPLICANT: Li, Li
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Mishra, Vishnu
 ; APPLICANT: Shenoy, Suresh G.
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Tchernev, Velizar T.
 ; APPLICANT: Vernet, Corine A.M.
 ; APPLICANT: Zerhusen, Bryan D.
 ; APPLICANT: Malyankar, Uriel M.
 ; APPLICANT: Guo, Xiaojia
 ; APPLICANT: Miller, Charles E.
 ; APPLICANT: Gangolli, Esha A.
 ; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 ; FILE REFERENCE: 21402-214 CIP
 ; CURRENT APPLICATION NUMBER: US/10/087,684
 ; CURRENT FILING DATE: 2003-03-10
 ; PRIOR APPLICATION NUMBER: 60/253,834
 ; PRIOR FILING DATE: 2000-11-29
 ; PRIOR APPLICATION NUMBER: 60/250,926
 ; PRIOR FILING DATE: 2000-11-30
 ; PRIOR APPLICATION NUMBER: 60/264,180
 ; PRIOR FILING DATE: 2001-01-25
 ; PRIOR APPLICATION NUMBER: 60/274,194
 ; PRIOR FILING DATE: 2001-03-08
 ; PRIOR APPLICATION NUMBER: 60/313,656
 ; PRIOR FILING DATE: 2001-08-20
 ; PRIOR APPLICATION NUMBER: 60/327,456
 ; PRIOR FILING DATE: 2001-10-05
 ; NUMBER OF SEQ ID NOS: 220
 ; SOFTWARE: CuiSeqlist version 0.1
 ; SEQ ID NO 1
 ; LENGTH: 2860
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (59)..(2857)
 US-10-087-684-1

Query Match 30.9%; Score 889; DB 13; Length 2860;
 Best Local Similarity 60.7%; Pred. No. 2,1e-211;
 Matches 1637; Conservative 0; Mismatches 960; Indels 102; Gaps 7;

QY 184 ACCGAGTGGTGCACACCGGACCTGTCCTCCACATTCCTGGGAGCCCGAGGATG 243
 Db 168 ATCTCCTCCCGTCAGCGCCAGCAGAGCCGCTGCTCTCTCTGAGGACCCACAGGACG 227
 QY 244 TGTATATCTCAAGAACAGCCAGTGTCTGTTGTGCGAAGCGCGTGGCCCGCAGCAGCA 303
 Db 228 CCTATATTGTGAAGAACAGCCCTGTGGAGCTCCGTGGCGCGCCTTCCCGCCACACAGA 287

QY 304 TCTTCTTCAAGTGCACCGGGAGTGGGTGCGCCAGGTGGACACACTGTATTCGAGCGCAGCA 363
 Db 288 TCTACTTCAAGTGCACCGGGAGTGGGTGCGCCAGGTGGACACACTGTATTCGAGCGCAGCA 347
 QY 364 CAGACGGGAGCAGTGGTGCAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCGACGAG 423
 Db 348 TGGATGAGGCCACCGGTCTGCGGGTGCAGAGGTGCAGATCGAGGTGTCGGGGCAGCAGG 407
 QY 424 TCGAAGAGGTCTTCGGGCTGGAGGAATACTGCTGCCAGTGGCTGGCATGGAGTCTCCTCGG 483
 Db 408 TGGAGGAGCTCTTTGGGCTGGAGGATTAATGCTGCCAGTGGCTGGCATGGAGTCTCCTCGG 467
 QY 484 GCACACCAAGAGTCAGAGGCCCTACATCCCATAGCCAGATTGCGCAAGAACTTCGAGC 543
 Db 468 GCACACCAAGAGTCGCGGAGCTACGTCGCGCATCGCTCTGCGCAAGAACTTCGATC 527
 QY 544 AGGAGCGCTGGCCCAAGGAGGTGCCCTGGAGAGGAGCATCGTGTGCTGCCCTGCCGTCAC 603
 Db 528 AGGAGCTCTGGGCAAGAGGTGCCCTGGAGCATGAGGTTCTCTGAGTGGCGCGCGC 587
 QY 604 CGGAGGCGATCCCTCCAGCCGAGGTGGGTCCGGAACGAGCATCTGGTGCACCGGT 663
 Db 588 CGGAGGCGGTGCTGTGCGGAGGTGGAATGGCTCAAGATGAGGATGTCTATCGACCCCA 647
 QY 664 CCCTGGACCCCAATGTATATATATCAGCGGAGCAGCAGCTGGTGGTGCAGACGCGCGC 723
 Db 648 CCCAGGACACCAACTTCTGCTCACCATCGACCAACACTCATCTCGCGCAGCGCGCGC 707
 QY 724 TTGCTGACACGGGCAACTACACCTGCTGGTGGCAGACATCGTGGCAGTGGCGGAGCG 783
 Db 708 TGTGCGACATCGCACTATACCTGCTGGTGGCAGAAACATCGTGGCGCAACGCGCGGAGCA 767
 QY 784 CCTCGCTGCTGTATCTGCTAGCTGAAACGCTGGTGGTGGTGCAGCTGACCGAGTGGTCCG 843
 Db 768 CCACTGCCACCGTCTGCTGCTAGTGAATGGCGGCTGGTCCAGCTGGGAGAGTGGTTCAC 827
 QY 844 TCTGCGAGCGCAGCTGTGGCGCGGCTGGCAGAAACGAGCCGAGCTGCACCAACCCCGG 903
 Db 828 CCTGCTCCAAACCGCTGTGGCGGAGGCTGGCAGAAAGCGCACCCGACCTGCACCAACCCCG 887
 QY 904 CGCTCTCTCAACGGGCGCTTTCTGTGAGGGGCGAGATGTCCATGACCGGACCGTCTCT 963
 Db 888 CTCACATCAACGAGGAGGCGCTTCTGCGAGGCGCAGGATTCAGAA--GACCGCTGCA 944
 QY 964 CTCTGCTGTCTCTGTGGAGCGGAGCTGGAGCCGCTGGAGCAAGTGGTGGCGCTGTGGG 1023
 Db 945 CCACCATCTGCCAGTCGATGGGCGTGGACGAGTGGAGCAAGTGGTTCAGCTTGAGCA 1004
 QY 1024 TGGACTGCACCCACTGGCGGAGCGGTGAGTGTCTGACCCAGCAGCCCGCAACGAGGGG 1083
 Db 1005 CTGAGTGTGCCACTGGCGTGGCGGAGTGCATGGCGGCCCGCCACCCAGAGCGGAGGCC 1064
 QY 1084 AGGAGTGCAGGGGCACTGACCTGGACACCCGCACTGTACCACTGACCTCTGTGTACACA 1143
 Db 1065 GTGACTGAGCGGAGCGCTGCTGCTTAAGAACTGCACAGATGGGCTGTGCATCAAC 1124
 QY 1144 GTGCTTCTGGCCCTGAGGAGCTGGGCGCTCTATGTGGGCGCTCATCGCGCTGGCGCTGCC 1203
 Db 1125 TGGAGGCTCTAGGGGATGCGGCGCTGTATGGCGGCGCTGTGGTGGCCATCTTCGTGGTGG 1184
 QY 1204 TGGTCTGCTGCTGTTGCTCTATCTCGTTTATTCGCGAAGAGGAGGGGCTGGACT 1263
 Db 1185 TGGCAATCTCTATGGCGGTGGGGGTGGTGTGTACCGCGCAACTGCGCTGACTTCGACA 1244
 QY 1264 CAGATGTGGCTGACTCGTCCATT---CTCAGCTCAGGCTTCAGCGCGCTGACATCAAGC 1320
 Db 1245 CAGACATCACTGACTCATCTGCTGCCCTGACTGGTGGTTCACCGCGCTCAACTTTAAGA 1304
 QY 1321 CCAGCAAGCAGACAAACCCCATCTGCT-----CACCATCCAGCGGAGCTCAGACCA 1374
 Db 1305 CGGCAAGGCCAGTAACCGCAGCTCTACACCCCTCTGTGCTCTGACCTGACGCA 1364
 QY 1375 CCACCACTACAGGCGAGTCTGTGTCGCCGCGAGGATGGGCGCCAGCCCCCAAGTTCCAGC 1434

Db 1365 GGCAGGATATACCGGAGACCGTGTATGCTGAGGACTCCACGACAAATCCCCA 1424
 QY 1435 TCACCAATGGGACCTGCTCAGCCCC-----1460
 Db 1425 TGACCACTCTCTCTGCTGGAGCCCTTACCAGCCTTAAGGTCAAGGTCTACAGTCCA 1484
 QY 1461 -----CTGGGTGGGGCGGCC 1476
 Db 1485 GCACCAAGGCTCTGGGCCAGGCTTGCGAGATGGGCTGACCTGCTGGGGGTCTTCCCGC 1544
 QY 1477 ACACACTGACACAGCTCTCCACCTCTGAGCGGAGGTCTCTCCCGCTCTCCA 1536
 Db 1545 CTGGACATACCTTAGGATTTCCGCCGGGACACCTCTCTCTGACCTGCGGAGGCCA 1604
 QY 1537 CCAGAACTAC-----TTCCGCTCTCCCTGCGCGAGSCACAGCAACATGACCTATGGGA 1590
 Db 1605 GCTCGGTTCCAGCAGCTCTTGCGCTCTGCGGAGACCCAGGAGCAGCGTCAAGGCCA 1664
 QY 1591 CTTCAACTCTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCTCTCTCATCC 1650
 Db 1665 CTTTGGCTCTCTGGGTGGGAGGCTCAGATCCCGGCAAGGGGTGAGCTTGCTGGTGC 1724
 QY 1651 CCCAGATGCATACCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGG 1710
 Db 1725 CCATGAGGCAATTCGCCAGGCAAGTTCTACGAGATGATCTACTCATCAAGGCCAG 1784
 QY 1711 AAGACGTGAGGTTGCCCTAGCTGGCTGTGAGCCCTGCTGAGTCCCATCGTTAGCTGTG 1770
 Db 1785 AAGTACCTCGCGCTTTAGAGGGAGCCAGACAGTATTGAGCCCTCGGTGACCTGTG 1844
 QY 1771 GACCCCTGGGCTCTGCTACCCGCGCAGTATCTCTGGCTATGACCTGTGGGAGC 1830
 Db 1845 GACCCACAGGCTCTGCTGTGCGCGCCGCTCATCTTCAATGCCCTGCTGCCAAG 1904
 QY 1831 CCAGCCCTGACAGCTGGAGCTCGCTCTAAAGCAGTCTGCGAGGCGAGCTGGGAGC 1890
 Db 1905 TCAGTCCCGTGAAGTATCTTTGAGTCAAGACCCAGGCCCCACAGGCGACCTGGAG- 1963
 QY 1891 AGGATGTCTGACCTGGCGAGGAGGCGCCCTCCACCTCTACTATGCGAGCTGGAGG 1950
 Db 1964 --GAGTGGTGACCTGGATGAGGAGACCTTGAACACACCTCTACTGCGAGCTGGAGC 2021
 QY 1951 CCAGTGCCTGCTAGCTCTTACCGAGCAGCTGGGCGGCTTTGCCCTGCTGGAGGAGGCC 2010
 Db 2022 CCAGGCGCTGTACATCTCTGCTGACACAGCTGGGCACTGCTGTTACGCGCGAGTCT 2081
 QY 2011 TCAGCTGGCTGCGCCAGCGCTCTCAAGCTGCTTCTGTTGCGCGGCTGGCTGCACCT 2070
 Db 2082 ATTCCGCTCAGAGTCAAGCGGCTCAGCTGGCGGCTTTCGCCCGCCCTCTGCACCT 2141
 QY 2071 CCCTCGAGTACAACATCCGGGTCTACTGCTGATGACACCCAGATGCACTCAAGGAGG 2130
 Db 2142 CCCTGGAGTACAGCTCTCCGGGTCTACTGCTGGAGACACGCTGTAGCAGTGAAGGAG 2201
 QY 2131 TGGTGCAGCTGAGAGCAGCTGGGGGACAGCTGATCCAGAGCCACGGGTCTGCACT 2190
 Db 2202 TGCTGGAGCTGGAGCGGACTCTGGGCGGATCTTGGTGGAGGAGCGGAAACCGCTAATGT 2261
 QY 2191 TCAAGCAGATTACACACACTGCGCTTATCCATCCAGGATGTGCCAGCTCCCTGTGGA 2250
 Db 2262 TCAAGCAGATTACACACACTGCGCTTATCCCTCATGACCTCCCGCCCATTTGGA 2321
 QY 2251 AGAGTAAGCTCTGCTGAGTACCAAGGAGATCCCTTTTATCAGATCTGGAATGCGAGC 2310
 Db 2322 GGAGCAAGCTGTGGCCAAATACCAAGAGATCCCTTCTATCAGATTTGGAGTGGCAGCC 2381
 QY 2311 ACCGTAATTGCACTGACCTTCACTGAGGCTGAGCGGCTGAGCCTAGTGAAGCTGG 2370
 Db 2382 AGAAGGCCCTTCACTGCACTTTACCTGAGGAGGACAGCTTGGCTCCACAGAGCTCA 2441
 QY 2371 CTTGCAAGCTGTGGGTGTGGCAGGTGGAGGCGAGCGGCGAGAGCTTCACTCACTCA 2430

Db 2442 CTTGCAAGATCTGCTGCGCAAGTGAAGGGAGGGCCAGATATTTCCAGCTGCATACCA 2501
 QY 2431 ACATCACCAAG--GACACAAAGTTTGTGAGTGTGCTGTGAGAGTGAAGCGGGG 2487
 Db 2502 CTTGCGAGAGACACTGCTGCTCCCTGGAACACTCTCTGCTCTGCGGAGCAGCTG 2561
 QY 2488 TCCAGCCCTGTTGGGCCCCAGTGCCTTCAAGATGCCCTTCTCTCATTTGGCAGAGATAA 2547
 Db 2562 TCACCAACCAGCTGGGACCTTATGCTTCAAGATCCCACTGTCTCATCCGCGCAGAATAT 2621
 QY 2548 TTTTCAGCTGAGACCCACCTCTAGCGGGGTGCGGAGTCTTGGGCCAGAAAC 2607
 Db 2622 GCAACAGCTAGATGCCCCCACTCACGGGGCAATGACTGGCGGATGTTAGCACAAGC 2681
 QY 2608 TCCACCTGAGACGCACTCTCAGCTTCTTTCCTTCAAGCCAGCCAGCCCATGATCC 2667
 Db 2682 TCTCTATGACCGGTACTGATTTACTTTGCGACCAAGAGGAGCCCGGCTGTATCC 2741
 QY 2668 TCAACCTGTGGGAGGCGGCACTTCCCCAAACCGGCAACTCAGCCAGCTGGCTGCAGCAG 2727
 Db 2742 TGGACCTCTGGGAAGCTCTGACGAGGACGATGGGACCTCAACAGCTGGCGAGTGCCT 2801
 QY 2728 TGGCTGAGCTGGGCGAGCAGACGCTGGCTCTTTCAGGTGTCGAGGCTGAGTGTGA 2786
 Db 2802 TGGAGGAGATGGGCAAGAGTGAATGCTGTGGTGTGGCCACCGACGCGGAGTGTCTGA 2860

RESULT 13

US-10-218-779-1
 ; Sequence 1, Application US/10218779
 ; Publication No. US2004002922A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Edinger, Shlomit
 ; APPLICANT: MacDougall, John
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Stone, David
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Grosse, William
 ; APPLICANT: Alsbrook II, John
 ; APPLICANT: Lepley, Denise
 ; APPLICANT: Rieger, Daniel
 ; APPLICANT: Burgess, Catherine
 ; APPLICANT: Casman, Stacie
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Boldog, Ferenc
 ; APPLICANT: Li, Li
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Mishra, Vishnu
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Shenoy, Suresh
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Tchernev, Velizar
 ; APPLICANT: Vernet, Corine
 ; APPLICANT: Zerhusen, Bryan
 ; APPLICANT: Malyankar, Uriel
 ; APPLICANT: Guo, Xiaojia
 ; APPLICANT: Miller, Charles
 ; APPLICANT: Gangolli, Esha
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-214
 ; CURRENT APPLICATION NUMBER: US/10/218,779
 ; CURRENT FILING DATE: 2002-08-14
 ; PRIOR APPLICATION NUMBER: 60/253,834
 ; PRIOR FILING DATE: 2000-11-29
 ; PRIOR APPLICATION NUMBER: 60/250,-926
 ; PRIOR FILING DATE: 2000-11-30
 ; PRIOR APPLICATION NUMBER: 60/264,180
 ; PRIOR FILING DATE: 2001-01-25
 ; PRIOR APPLICATION NUMBER: 60/313,656
 ; PRIOR FILING DATE: 2001-08-20
 ; PRIOR APPLICATION NUMBER: 60/327,456
 ; PRIOR FILING DATE: 2001-10-05

; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-218-779-1

Query Match 30.9%; Score 889; DB 13; Length 2860;
Best Local Similarity 60.7%; Pred. No. 2,1e-211;
Matches 1637; Conservative 0; Mismatches 960; Indels 102; Gaps 7;

```
QY 184 ACCAGTGCCTGGTGCACCGGACCTGCTTCCCACTTCTGTGTGGAGCCGAGGATG 243
Db 168 ACTCCTTCCCTCGCGCCAGCAGAGCCGCTGCCCTACTTCTGCGAGGACACAGGAGC 227
QY 244 TGTATATCGTCAAGAACAGCAGCTGCTGTGTGCAAGGCGGTGCGCGCAGCGAGA 303
Db 228 CTTACATTGTGAAGAACAGCCTGTGGAGCTCCGCTGCGCGGCCCTTCCCGCCACAGA 287
QY 304 TCTTCTTCAAGTGCAACCGGGAGTGGGTGCGCCAGGTGGACCAAGTGTGATCGAGCGCAGCA 363
Db 288 TCTACTTCAAGTGTCAACCGCGAGTGGGTGACCGAAGACACAGTCAACAGGAAGGCC 347
QY 364 CAGAGCGGAGCAGTGGTGAAGCCGACCAATGGAGGTCCGCAATTAATGTCTCAAGGCGAGG 423
Db 348 TGGATGAGGCGACCGGCTCTGCGGGTGGCGAGGTGCAGATCGAGGTGTGCGGCGAGCAGG 407
QY 424 TCGAAGAGGTGTTTCGGGCTGGAGGAATACTGTTCCAGTGGGTGGCATGGAGCTTCCTGG 483
Db 408 TGGAGGAGCTCTTTGGGCTGGAGGATTAATGGTGCAGTGGGTGGGTGGAGCTTCCGCGG 467
QY 484 GCACCAACAGAGTCAAGAGCCTTACATCCGATAGCCAGATTGCGCAAGAACTTTCGAGC 543
Db 468 GCACCAACAGAGTCCGCGAGCTTACGTGCGCATCGCTACCTCGGCAAGAACTTTCGATC 527
QY 544 AGAGCGCGTGGCCAAAGAGGTGTCCTGGAGCAGGCGATGCTGCTCCCTGCGCGTCCAC 603
Db 528 AGGAGCCTCTGGGCAAGAGGTGCCCTGGACCATGAGGTTCCTTCGAGTGGCGCCCGC 587
QY 604 CGGAGGCGATCCCTCCAGCGGAGTGGAGTGGCTTCGGAAGAGGACCTGGTGGACCCGT 663
Db 588 CGGAGGCGGTGCTGTGGCGGAGTGGAGTGGATGGGTCAAGAAATGAGGATGTATCGACCCCA 647
QY 664 CCTTGAACCCCAATGTATATACATCAGCGGAGCAGACGCTGCTGGTGGCAGCGTCCGCGAGCG 723
Db 648 CCCAGGACACCAACTTCTGCTCACCATCAGCACACCTCATCATCCGCGAGGCGCGCC 707
QY 724 TTGCTGACACGGCCAACTACCTGCGTGGCCCAAGAAACATCGTGGCAGCTGCGCGAGCG 783
Db 708 TGTGAGACATGCGCAACTATACCTGCGTGGCCCAAGAAACATCGTGGCCAAACGCGGAGCA 767
QY 784 CTTCCGCTGTGTCATCGTCTAGCTGAACGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 843
Db 768 CCACTGCCACCGTATCTGCTAGTGAATGGCGGCTGGTCCAGCTGGGCGAGAGTGGTTCAC 827
QY 844 TCTGACGCGCCAGCTGTGGCGCGGCTGGCAGAAACGAGCGCGGAGCTGACCAACCCCG 903
Db 828 CTTGCTCCAAACCGCTGTGGCGGAGGTGGCAGAGCGCACCCGAGCTGACCAACCCCG 887
QY 904 CGGCTCTCAACGGGGCGCTTTCTGTGAGGGGCGAATGTTCATGACCGGACCGCTCTCT 963
Db 888 CTTCACTCAACGAGGGGCGCTTCTGCGAGGGCCAGGCAATTCAGAA--GACGCGCTGCA 944
QY 964 CTTGCTGTCTCTGTGACCGGAGCTGGAGCCCGTGGAGCAAGTGGTGGGCTGTGGGCG 1023
Db 945 CCACCAATCTGCGCAGTGCATGGGCGGTGGACGAGTGGAGCAAGTGGTCAAGCTGAGCGCA 1004
QY 1024 TGGACTGCAACCACTGGCGGAGCGGTGAGTGTCTTGACCCAGCACCCGCAACGAGGGG 1083
Db 1005 CTGAGTGTGCCCACTGGCGGTGAGTGTGATGGCGGCCCAACCCCAAGACGAGGCGC 1064
QY 1084 AGGAGTGCCAGGGCACTGACTGGACACCCGCAACTGTGTACAGTGCACCTCTGTGTACACA 1143
```

```
Db 1065 GTGACTGCAGCGGAGCGTGTGTGACTTAAAGACTGCACAGATGGGTGTGCATGCAAC 1124
QY 1144 GTGCTTCTGGCCCTCAGGACGTGGCCCTCTATGTGGGCTCATCGCCGTGGGCGCTGCC 1203
Db 1125 TGGAGGCTCAGGGGATGCGGCGTGTATGCGGGCTGTGGTGGCCATCTTCGTGGTGG 1184
QY 1204 TGTCTCTGCTGCTTGTCTCTCATCTCTGTTATTTGCGGGAAGAGGAGGCGCTGGACT 1263
Db 1185 TGGCAATCTCATGCGGCTGGGGTGGTGTACCGCGCAACTGCGCTGACTTCGACTGACA 1244
QY 1264 CAGATGTGGCTGACTCTGCTCAATT---CTCAGCTCAGGCTTCAGCCCGTTCAGACTCAAG 1320
Db 1245 CAGATCAGCTGACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1304
QY 1321 CCAGAAAGCAGACAAACCCCATCTGCT-----CACCATCCAGCGGACCTCAGCACCA 1374
Db 1305 CGGCAAGGCCAGTAAACCGCAGCTCTACACCCCTCTGTGCTGCTGCTGCTGCTGCTGCTG 1364
QY 1375 CCACCACTTACAGGGCAGTCTCTGTCCCGGAGGATGGGCGCCAGCCCAAGTTCCAGC 1434
Db 1365 GCGCGCGCATCTACCGCGGAGCCGCTGTATGCTGCTGAGGACTCCACCGACAAATCCCCA 1424
QY 1435 TCACCAATGGCACCTGCTCAGCCCC-----TGAGCCTTAAAGCTTACAGCTCCA 1460
Db 1425 TGACCAACTCTCTCTGCTGAGCCCCCTTACCGAGCTTAAAGCTTACAGCTCCA 1484
QY 1461 -----CTGGTGGCGGCGCC 1476
Db 1485 GCACCAAGGCTCTGGCCAGGCTGCGCAGATGGGGCTGACCTGCTGGGGTCTTGGCGC 1544
QY 1477 ACACATGCAACAGCTCTCCCACTCTGAGGCCGAGGAGTTCGTCTCCGCTCTCTCCA 1536
Db 1545 CTGSCACATACCTTAGCGAATTCGCGCGGACACCCACTTCTCTGCACTGCGCAGCGCCA 1604
QY 1537 CCAGAAACTAC-----TTCGGCTCCCTGCGCGGAGGCAACAGCAATGACCTATGGGA 1590
Db 1605 GCTTCGGTTCACAGCAGCTCTTGGGCGCTGCGCGGAGACCCAGGAGCAGCGTCAGCGCA 1664
QY 1591 CTTTCAACTTCTTCGGGGCGGCTGATGATCCCTAATACAGTATACAGCTTACAGCTCTCTCATCC 1650
Db 1665 CTTTGGCTGCTTGGTGGGAGGCTCAGCATCCCGGCAAGGGGTGAGTTCGTGGTGGTGC 1724
QY 1651 CCCAGATGCATACCCCGAGGGAAGATCTAGATCTACCTCAGCTGCAACAAGCGCG 1710
Db 1725 CCAATGGAGCAATCCCGAGGCAAGTTCACAGATGATCTACTCATCAACAGGCGAG 1784
QY 1711 AAGACGTGAGGTTGCCCTTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCTGTTAGCTGTG 1770
Db 1785 AAGTACCTGCGCTTTCAGAAAGGAGCCACAGATATTGAGCCCCCTCGGTGACCTGTG 1844
QY 1771 GACCCCTGCGGCTCTGCTCACCCGCGCAGTCACTCTGGCTATGGACCACTGTGGGAGC 1830
Db 1845 GACCCACAGGCTCTCTGCTGTGCGGCGCGCTCATCTCACCATGCCCCACTGTGCCAAG 1904
QY 1831 CAGCCCTGACAGCTGGAGCTGCGCTCAAAGACAGTCTGTCGAGGAGCAGCTGGGAGC 1890
Db 1905 TCAGTGGCGGCTGACTGGATCTTTCAGCTCAAGACCCAGGCCACAGGSCCATCTGGAG- 1963
QY 1891 AGGATGTGCTGACCTTGGCGGAGGAGCGCCCTCCACCTCTACTACTGCTGCTGAGGAGG 1950
Db 1964 --GAGTGTGACCTTGGATGAGGAGCCCTGAACACACCCCTGCTACTGCCAGTGGAGC 2021
QY 1951 CCAGTGCCTGCTACGCTCTTCAACCGCAGCTGGCGGCTTTCCCTGCTGGGAGGAGGCC 2010
Db 2022 CCAGGCGCTGTACATCTCTGCTGGACAGCTGGGCACTTACGTTTACGGGCGAGTCTCT 2081
QY 2011 TCAGCGTGGCTGCGCCAAAGCGCTCTAAGCTCTTCTGTTTTCGCGGCTGGCTGCACT 2070
Db 2082 ATTCCCGCTCAGCAGTCAAGCGGCTCCAGTGGCGCTCTTCGCCCCCGCCCTCTGCACT 2141
QY 2071 CCCTCGAGTACAACTCCGGGTCTACTGCTGCTGATGACACCCAGCATGCACTCAAGGAGG 2130
```

2142	CCCTGGAGTACAGCCTCCGGGTTACTGTGCTGGAGGACACGCTGTAGCACTGAGAGAGG	2201
2131	TGCTGCAGCTGGAGAAAGACAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCTCTCACT	2190
2202	TGCTGGAGCTGGAGCGGACTCTGGCGGATCTTGGTGGAGGAGCGAAACCGCTAATGT	2261
2191	TCAAGGACAGTTTACCACAACCTGGGCTATTCCATCCACGATGTGCCAGCTCCCTGTGGA	2250
2262	TCAAGGACAGTTTACCACAACCTGGGCTCTCCCTCCATGACCTCCGCCATGGCCCATTGG	2321
2251	AGAGTAGCTCCCTGTGTGAGTACCAAGAGATCCCTTTTATCATCTGGAATGGCAGCG	2310
2322	GGAGCAAGCTGTGGCCCAATATCCAGAGATCCCTTCTATCATCTTGGATGGCAGCC	2381
2311	AGCGGTACTTGCATGCACTTCACTCTGGAGCGGTGTACGCCCGACACTAGTGTGACTGG	2370
2382	AGAAAGCCCTCCACTGCACTTTCACTCTGGAGAGGACAGCTTGGCTCCACAGAGCTCA	2441
2371	CCTGCAAGCTGTGGGTGGCAGGTGAGGGCGGACGGGAGAGCTTCAGCATCAACTTCA	2430
2442	CCTGCAAGATCTGGTGGCGCAAGTGGAAAGGGAGGGCCAGATATTCAGCTGCATACCA	2501
2431	ACATCACCAAG--GACACAAAGTTTGTGAGCTGTGCTGGCTCTGGAGAGTGAAGCGGGG	2487
2502	CTCTGGCAGAGACACCTCTGCTCCCTGGACACTCTCTGCTGTGCCCTGGCAGCACTG	2561
2488	TCCAGAGCCTGGTGGGCCCCAGTGGCTTCAAGATCCCTTCTCTCATTTGGGACAGATAA	2547
2562	TCAGCACCGAGCTGGGACCTTATGCTTTCAGATGCCACTGTCCATCCGCCACAGAAATAT	2621
2548	TTTCCAGCTGAGCCACCCCTGTAGGGGGGGTGGCGACTGGCGGACTCTGGCCCCAGAAAC	2607
2622	GCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGATGTTAGCACAGAAGC	2681
2608	TCCACCTGGACAGCCATCTCAGCTTCTTTGGCTTCCAAAGCCCGAGCCCCACAGCCCATGCC	2667
2682	TCTCTATGGACCGGTACCTGAATTACTTTTGGCAGCAAAAGCGAGCCCCACGGGTGTGATCC	2741
2668	TCAACCTGTGGAGGGCGGCACTTCCCACAGCGCAACCTCAGCCAGCTGGCTCAGCAG	2727
2742	TGGACCTCTGGGAAGCTCTGACAGAGCAGATGGGGACCTCAACAGCCTGGCGAGTGCCT	2801
2728	TGGCTGGAGCTGGGCGAGCCAGACGCTGGCCCTTTTACAGTGTCCGAGGCTCAGTGCTCA	2786
2802	TGGAGCATGATGGCAAGATGTGAGATGCTGGTGGTGTCTGCGCCACCGAGGGGACTGCTGA	2860

RESULT 14

US-10-087-684.-3
; Sequence 3, Application US/10087684
; Publication NO. US20040029116A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: MacDougall, John R.
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David J.
; APPLICANT: Grosse, William M.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie, J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zethusen, Bryan D.
; APPLICANT: Malyankar, Uriel M.

QY	784	CTCTCGCTGCTGTCTATCGTCTACGTCAACCGGTGGGTGGTGCAGCTGGACCGAGTGGTCCG	843
Db	768	CCACTGCCACCGTCAATCGTCTACGTGAATGGCGGCTGTTCCAGCTGGCGAGAGTGGTCAAC	827
QY	844	TTCTGACGCCAGCTGTGTGGCCGCGCTGGCGAAGAACGAGCGGAGTGTGCACCAACCGGG	903
Db	828	CGTGTCCAACCGCTGTGGCCGAGGCTGGCAAGCGCACCGGACCTGCACCAACCGCC	887
QY	904	CGCCTCTCAACGGGGCGCTTTCTCTGAGGGGCAAGATGTCCATGACCGCACCGTCTCCT	963
Db	888	CTTCACTCAACGAGGGGCTTTCTCGAGGGCCAGGCAATCCAGAA---GACCGCGTGA	944
QY	964	CTCTCTTTGTCTGTGTGAACGCGACCTGGAGCCCGCTGAGCAAGTGGTGGCCGTGGGC	1023
Db	945	CCACCATCTGCCCACTGCATGGGGCTGGACGGAGTGAGGCAAGTGGTACGCTTGACGA	1004
QY	1024	TGGACTGACCCACTGGCCGAGCCGTGAGTGTCTGACCCAGCACCCCGCAACGAGGGGG	1083
Db	1005	CTGAGTGTGCCCACTGGCGGTAGCCGCGAGTCATGGCGCCCGCCACCCAGAACGAGGCC	1064
QY	1084	AGGAGTGCACGGCACTGACCTGGACACCCGCAACTGTACAGTGAACCTCTGTGTACACA	1143
Db	1065	GTGACTGACGGGAGCGTGTCTGCACTTAAGAACTGCACAGATGGGCTGTGATGCAAC	1124
QY	1144	GTGCTTTGGCCCTGAGACAGTGGCCCTCTATGTGGGCTCATCGCCGTGGCGTCTGCC	1203
Db	1125	TGGAGGCTTCAGGGGATGCGCGCTGTATGCGGGCTCTGGTGGCCATCTTGTGGTGC	1184
QY	1204	TGCTCTGCTGCTGTCTCTCATCTCTGTTTATGCGGAAAGAGGAGGGGTGTGACT	1263
Db	1185	TGGCAATCCTCATGCGGTGGGGGTGGTGGTGTACCGCGCACTGCGCGTCTTGACA	1244
QY	1264	CAGATGTGGTGACTCTGTCCATT---CTCACTCAGGCTTCCAGCCCGTCAAGCATCAGC	1320
Db	1245	CAGACATCACTGACTCATCTGTGCCCTGACTGGTGTGTTTCCACCCCGTCAACTTTAAGA	1304
QY	1321	CCAGCAAGCAGACAACCCCATCTGCT-----CACCATCCAGCGGACCTCAGACCA	1374
Db	1305	CGGCAAGGCCAGTAACCCGAGTCTCTACCCCTCTGTGCCTCTGACTCTGACGCCA	1364
QY	1375	CCACCCTTACAGGGGAGTCTCTGTCCCGGAGGATGGGCCAGCCCAAGTTCAGC	1434
Db	1365	GGCGCGCATCTACCGCGGACCGGTATGCCCTGAGGACTCCACGACAAATCCCA	1424
QY	1435	TCACCAATGGCACTTGCTCAGCCCC-----	1460
Db	1425	TGACCAACTCTCTCTGTGGACCCCTTACCAGCCTTAAGGTCAAGGTCTACAGTCCA	1484
QY	1461	-----	1476
Db	1485	GCACCACGGGCTCTGGGCCAGGCTTGGCAGATGGGCTGACCTGTGGGGTCTTGCCGC	1544
QY	1477	ACACATGACACAGCTCTCCCACTCTGTAGGCCGAGGATTCGTCTCCCGCTCTCA	1536
Db	1545	CTGGCACAATACCCCTAGCGATTTCCGCCGGGACACCCACTTCCTGCACTTGCAGCGCA	1604
QY	1537	CCCAGAATCTAC-----TTCCGCTCTCCCGGAGGACACAGCAATGACCTATGGGA	1590
Db	1605	GCCTCGGTTCCAGCAGCTCTTGGGCTCTGCCCGAGACCCAGGAGCAGCTCAGCGCA	1664
QY	1591	CCTTCAACTTCTCGGGGGCGGCTGATGATCCCTAATACAGTATCAGCCTCTCATCC	1650
Db	1665	CTTTGGCTGCCTGGTGGAGGCTCAGCATCCCCGGCACAGGGTCACTGTTGTTGGTGC	1724
QY	1651	CCCCAGATGCCATACCCCGAGGGAAGATATGAGATCTACCTCAGCTGTGCAAGACCGG	1710
Db	1725	CCAATGGAGCAATCCCGAGGCAAGTTCTACGAGATGATCTACTCATCAACAGGAG	1784
QY	1711	AAGACGTGAGTTGCCCTAGCTGGCTGTGCAGACCCTGCTAGTCCCATCGTTAGCTGTG	1770
Db	1785	AAAGTACCTTGCCTTTTCAAGAGGACCCAGACAGTATTGAGCCCTCGTGAACCTGTG	1844

QY	1771	GA	CCCTCTGGCGTCTCTGCTACACCGGCGAGTCACTCTGGCTATGAGCACTGTGGGAGC	1833
Db	1845	GA	CCACACAGCGCTCTGCTGTGCGCGCCCGTCACTCTCACCATGCCCACTGTGCGCAG	1904
QY	1831	CC	AGCCCTTGACAGCTGGAGCCTGGCGCTCAAAAGACAGTCTGTGCGAGGCGAGCTGGAGC	1890
Db	1905	TC	AGTGCCCGTGACTGAGTCTTTTCTTCTAAGACCCAGGGCCCAACAGGGCCACTGGGAG-	1963
QY	1891	AG	ATGTCTGCACTCTGGCGAGGAGCGCCCTCCCACTCTACTACTGCCAGCTGGAGG	1950
Db	1964	--	GAGGTGGTACCTGGATGAGGAGACCCGTGAACACACCCCTGCTACTGCCAGCTGGAGC	2021
QY	1951	CC	AGTGCCTCTACTGCTTTCACGAGCAGCTGGGCGCTTTGGCCCTGGTGGAGAGGCC	2010
Db	2022	CA	GGCCCTGTCACTCTGCTGGACAGCTGGGCACTTACGTGTTTCACTGGCGAGTCCCT	2081
QY	2011	TC	AGCGTGGTCCGCCCAAGCGCTCAAGCTGTTCTGTGTTTGGCGCGTGGCCCTGCACCT	2070
Db	2082	AT	TCCGCTCAGCAGTCAAGCGCTCCAGTGGCCGCTTTCGCCCCCGCCCTCTGCACCT	2141
QY	2071	CC	CTGAGTACCAATCCGGGTCTATGCTGATGACACCCACGATGCACTCAAGGAGG	2130
Db	2142	CC	CTGGAGTACAGCTCCGGGTCTACTGCTGTGAGGACACGCTGTAGCACTGAAGGAGG	2201
QY	2131	TG	TGCAGCTTGGAGAACAGCTGGGGGACAGCTGATCCAGAGACCAAGGTCCTGCAT	2190
Db	2202	TG	CTGAGCTTGGAGCGGACTCTGGCGGATCTTGGTGAGGAGCGCAACCGCTAATGT	2261
QY	2191	TC	AGGACATTTACCAACCTCGCTATCCATCCAGATGTGCCAGCTCCCTGTGGA	2250
Db	2262	TC	AGGACATTTACCAACCTCGCTCTCCCTCCATGACCTCCCCATGCCCATTTGA	2321
QY	2251	AG	ATAGCTCCTTGTTCAGCTACAGAGATCCCTTTTATCACTCTGGAATGGCAGC	2310
Db	2322	GG	ACAGCTGTGGCCAAATACACGAGATCCCTTCTATCACATTTGGAGTGGCAGC	2381
QY	2311	AG	CGGTACTTGGCACTGCACCTTCACCTCGAGCGTCTCAGCCCGCAGCACTAGTGACCTGG	2370
Db	2382	AGA	GGCCCTCCACTTTCACCTTGAGAGGCACAGCTTGGCCTCCACAGAGCTCA	2441
QY	2371	CT	GTCAAGCTGTGGCAGTGGAGGCGCAGCGGACAGCTTCAGCATCAACTTCA	2430
Db	2442	CT	GTCAAGATCTCGTGTGGCAAGTGGAGGGAGGGCCAGATATTCAGAGCTGCATCCA	2501
QY	2431	AC	ATCACCAG---GACACAAGTTTCTGAGCTGTGGCTCTGAGAGTGAAGCGGGG	2487
Db	2502	CT	CTGGAGAGACACCTGTGGCTCCCTGGACACTCTGTCTCTGCCCTGGCAGCACTG	2561
QY	2488	TCC	AGCCCTGTGGGCCCCAGTGCCTTCAAGATCCCTTTCCTCATTCGGCAGAGATAA	2547
Db	2562	TC	ACCAACCCAGCTGGGACCTTATGCTTCAAGATCCCACTGTCCATCCGCAGAGATAT	2621
QY	2548	TT	TCCAGCTTGACCCACCTGTAGCGGGGTGCCACTGGCGGACTCTGCCCGAAGAAC	2607
Db	2622	GC	AAACCTTAGATGCCCCAACTCACGGGGCAATGATGGCGGATGTTAGCACAGAAGC	2681
QY	2608	TCC	ACCTGGACCACTTCAGTTCTTTTGCTCCCAAGCCAGCCCAACCAACCCATGATCC	2667
Db	2682	TCT	TATGACCGGTACCTGATTAATTGTCACCAAGAGCAGGCCCAACGGGTGTGATCC	2741
QY	2668	TC	AACTGTGGAGGCGGCACTTCCCCAAACGGCAACTTCAGCGAGCTGTGCTGCAGAG	2727
Db	2742	TG	AGCTCTGGGAAGCTCTGCAGCAGGACGATGGGACCTCAACAGCTGCGGAGTGCCT	2801
QY	2728	TGG	CTGGACTGGGCCAGCAGACGCTGGCCTTTCACAGTGTCCGAGGCTGAGTGCTGA	2786
Db	2802	TG	GAGATGGGCAAGAGTGAATGCTGTGGTGTGGCCACCGACGGGACTGCTGA	2860

GENERAL INFORMATION:		US-10-218-779-3	
APPLICANT: Edinger, Shlomit		Query Match	30.8%; Score 887.4; DB 13; Length 2860;
APPLICANT: MacDougall, John		Best Local Similarity	60.6%; Pred. No. 5.4e-211;
APPLICANT: Millet, Isabelle		Matches 1636; Conservative	0; Mismatches 961; Indels 102; Gaps 7;
APPLICANT: Ellerman, Karen			
APPLICANT: Stone, David		QY 184	ACCCAGTCCCTGGTCCCAACCGGACCTGCTTCCCACTTCTGTTGGAGCCCGAGGATG 243
APPLICANT: Gerlach, Valerie		DB 168	ACTCCTCCCGTACGCGAGCAGAGCCGCTGCTCCCTACTTCTTCTGAGAGCCACAGGACG 227
APPLICANT: Grosse, William		QY 244	TGTACATCGTCAAGAACCAAGCCAGTGTCTTGTGTGCAAGCCCGTGCCTCCGACGACG 303
APPLICANT: Alsobrook II, John		DB 228	CCTACATTTGTGAGAACCAAGCCCTGTGGAGCTTCTGCTGCGCGCTTCCCGCCACACAGA 287
APPLICANT: Lepley, Denise		QY 304	TCTTCTTCAAGTGAACCGGGAGTGGGTGCGGCGAGGTGACCACTGATCGAGCCGACGA 363
APPLICANT: Rieger, Daniel		DB 288	TCTACTTCAAGTGAACCGGGAGTGGGTGAGCCAGAACACACCTGTCACAGGAAGGCC 347
APPLICANT: Burgess, Catherine		QY 364	CAGAGCGGAGAGTGTGAGCGGACCATGGAGTCCGATTAAATGTCTCAAGGACGACG 423
APPLICANT: Casman, Stacie		DB 348	TGGATGAGCCACCGCCCTGCGGTGCGGAGGTGAGATCAGGTGTGCGGAGCAGG 407
APPLICANT: Spytek, Kimberly		QY 424	TCCAGAAAGTGTTCGGGCTGGAGGAATACTGTGTGCGAGTGCCTGGGATGAGCTCTCTGG 483
APPLICANT: Boldog, Ferenc		DB 408	TGGAGGAGCTCTTTGGGCTGGAGGATTACTGTGTGCGAGTGCCTGGGCTGGGCTCCGACG 467
APPLICANT: Li, Li			
APPLICANT: Padigaru, Muralidhara			
APPLICANT: Mishra, Vishnu			
APPLICANT: Patturajan, Meera			
APPLICANT: Shenoy, Suresh			
APPLICANT: Rastelli, Luca			
APPLICANT: Tchernev, Velizar			
APPLICANT: Vernet, Corine			
APPLICANT: Zerhusen, Bryan			
APPLICANT: Malyankar, Uriel			
APPLICANT: Guo, Xiaojia			
APPLICANT: Miller, Charles			
APPLICANT: Gangolli, Esba			
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same			
FILE REFERENCE: 21402-214			
CURRENT APPLICATION NUMBER: US/10/218,779			
CURRENT FILING DATE: 2002-08-14			
PRIOR APPLICATION NUMBER: 60/253,834			
PRIOR FILING DATE: 2000-11-29			
PRIOR APPLICATION NUMBER: 60/250, -926			
PRIOR FILING DATE: 2000-11-30			
PRIOR APPLICATION NUMBER: 60/264,180			
PRIOR FILING DATE: 2001-01-25			
PRIOR APPLICATION NUMBER: 60/313,656			
PRIOR FILING DATE: 2001-08-20			
PRIOR APPLICATION NUMBER: 60/327,456			
PRIOR FILING DATE: 2001-10-05			
NUMBER OF SEQ ID NOS: 216			
SOFTWARE: Patent In Ver. 2.1			
SEQ ID NO 3			
LENGTH: 2860			
TYPE: DNA			
ORGANISM: Homo sapiens			
US-10-218-779-3			
QY	484	GCACCAACAGAGTCAAGAGGCTACATCGCATAGCCAGATTCGCCAAGAACTTCGAGC	543
DB	468	GCACCAACAGAGTCAAGAGGCTACATCGCATAGCCAGATTCGCCAAGAACTTCGATC	527
QY	544	AGGAGCCGCTGGCCAAAGGAGGTTCCTGGAGAGGGCATTCGTGCTGCGCTTCGCTCCAC	603
DB	528	AGGAGCCCTGTGGCAAGGAGGTGCCCTCGACCATGAGGTTCCTTCGATGCGCCGCG	587
QY	604	CGAGGGCATCCCTCAGCCGAGGTGGGTCCGAGACGAGGACCTGGTGGACCCGT	663
DB	588	CGAGGGGTGCTGCTGTGGCCGAGGTGGATTCGATCAAGATGAGGATGTCATGACCCCA	647
QY	664	CCCTGACCCCAATGTATATACATCACCGGAGGACACAGCCCTGGTGGTGGAGCCCGCC	723
DB	648	CCAGGACACCAACTTCCTGCTCACCATCGACCAACCTCATCATCGCCAGGCCCGCC	707
QY	724	TTGCTGACCGGCAACTACCTCGGTGGCCAAAGAACATCGTGGGACGCTGCGGAGCG	783
DB	708	TGTGGGACACTGGCAACTATACCTGGTGGCCAAAGAACATCGTGGCCAAAGCGGGAGCA	767
QY	784	CTTCCGCTGTCTCATGCTAGTGAACGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	843
DB	768	CCATGCGCCGCTCATGCTTACGTAATGGCGCTGCTCCAGCTGGGAGAGTGGTTCAC	827
QY	844	TTGTCAGCGCCAGCTGTGGCGGGCTGGCAGAAACCGGAGCCGAGCTGCACCAACCCGG	903
DB	828	CTGTCTCAACCGCTGTGGCGGAGGTGGCAGAGCGCACCCGACCTGCACCAACCCCG	887
QY	904	CGCCTCTCAAAGGGGGCGCTTTCTGTAGGGGAGAAATGTCATGACCGACCGACCTCTCT	963
DB	888	CTCCACTCAACGGAGGGCGCTTCTGCGAGGGCCAGGCAATTCAGAA---GACCGCTGCA	944
QY	964	CTGTGCTTGTCTGTGTGAGCGGACGTGGAGCCCGTGGAGCAAGTGTGGCGCTGTGGGC	1023
DB	948	CCACCATCTGCCAGTCTGATGGGGCGGTGGAGCGAGTGGAGCAAGTGTGACCTGCGACA	1004
QY	1024	TGGACTGCACCCACTGGCGGAGCGGTGAGTGTCTGTGACCCAGCACCCCGCAACCGAGGG	1083
DB	1005	CTGAGTGTGCCCATCTGGCGTAGCCGCGAGTGTACCGCGCCACCCAGAACCGAGGCC	1064
QY	1084	AGGAGTGCAGGGCACTGACCTGACACCCCGCAACTGTACAGTACCTCTGTGTACACA	1143
DB	1065	GTGACTGCAGCGGACGCTGTCTGCACTTAAGAACTGCACAGATGGGTGTGCAATGCAAC	1124
QY	1144	GTGCTTTCGGCCCTGAGGACGTGGCCCTCTATGTGGGCTCATCGCGTGGCGCTGCGC	1203
DB	1125	TGGAGGCTTCAGGGGATGCGCGCGTGTATGCGGGGCTGCTGGTGGCCATCTTCGTGCTG	1184
QY	1204	TGGTCTGCTGCTGTCT	1263
DB	1185	TGGCAATCCTCATGGCGTGGGGTGGTGTGTACCGCGCAACTGCGGTGACTTCGACA	1244
QY	1264	CAGATGTGGTGTGCTGTGTCATT---CTCAGCTAGGCTTCAGGCGCTTCAGCATCAGC	1320
DB	1245	CAGACATCACTGACTCATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1304
QY	1321	CCAGCAAGCAGACCAACCCCATCTGCT-----CACCATCCAGCGGACCTTCAGACCA	1374
DB	1305	CGGCAAGCCGAGTAACCCGAGCTCTTACACCCCTCTGTGCTCTGCTGCTGCTGCTGCTG	1364
QY	1375	CCACCACCTACAGGGCAGTCTCTGTCCCGGAGGATGGGCGCCAGCCCGCCAGGTTCCAGC	1434
DB	1365	GGCGCGCATCTACCGCGGACCCGCTGTATGCCCTGCTGAGGACTTCCACCCGCAAAATCC	1424
QY	1435	TCACCAATGGGCACTGCTCAGCCCC-----	1460
DB	1425	TGACCAACTCTCTCTGCTGGACCCCTTACCCAGCCTTAAGTCAAGGTCTACAGCTCCA	1484
QY	1461	-----	1496
DB	1485	GCACCAACGGGCTCTGGGCGAGCCCTGGCAGATGGGGCTGACCTGCTGGGGCTTGGCGC	1544

QY	1477	ACACACTG	GACACCA	CAGCTCTCC	CACTCTG	AGGCGG	CAGGAGTT	CGTCTCC	CGCGCTCTCC	CA	1533							
DB	1545	CTGGCACA	TATCCCTAG	CGAATTTG	CGCGG	GACACCC	ACTTCT	CTGC	ACTCTGG	CAGCGCCA	1604							
QY	1537	CCCAGAA	CTAC-----	TTCCG	CTCCCTG	CCCGAG	GCACAG	CAAT	GACCTAT	GGGA	1590							
DB	1605	GCCTCG	GTTC	CAGAGCTCT	TTGG	CCCTG	CCCGAG	CCCGAG	AGGAGT	CAGGCTCAG	GGCA	1664						
QY	1591	CTTCAAC	TTCTCT	CGGGG	CCGGTGT	ATGATCC	CTTA	TACAG	GTATCAG	CTCTCAT	CC	1650						
DB	1665	CTTTTGG	TGCTCG	GTGG	AGGCTC	AGCATCC	CCG	GCACAG	GGT	CAGTTT	CTGCTGGTGC	1724						
QY	1651	CCCAGAT	GCATAC	CCCCCG	AGGAGATCT	ATAGATCT	ACTT	CAC	CTGCAT	CGCTGC	CAAGCCGG	1710						
DB	1725	CCAAT	GGAGCC	ATTCCC	CAGG	CAAGTTCT	AC	GAGATGT	ATCTACT	CA	CAAGG	CAG	1784					
QY	1711	AAGAC	GTGAG	GTG	CGCT	AGCTG	CTCAG	ACCTG	CTGCTAG	TCCCAT	GTGTAG	CTGTG	1770					
DB	1785	AAAGTA	CCCTGCC	GTTC	AG	AGG	ACCC	CAG	AGTATT	GAC	CCCTCG	TGACCTGTG	1844					
QY	1771	GACCC	CTTGG	CGTCT	GTCTCAC	CCG	CCAGTCT	CTTGG	CTATG	AG	CACTGTGG	GGGAGC	1830					
DB	1845	GACCA	CAGG	CTCTCT	GTG	CGCC	CGCTCAT	CTC	CA	CAATGCC	CACTGTG	CCGAAG	1904					
QY	1831	CCAGC	CTTGA	CAGTGG	AGCTG	CGCTCA	AAAG	CAGTGTG	TCG	AGG	GCAG	CTGGGAGC	1890					
DB	1905	TCAGT	GCCGTG	ACTGTG	ATCTT	CAGCT	CAAG	ACCC	AGG	CCCA	CTG	GGGAG-	1963					
QY	1891	AGGAT	GTGCTG	CACTGG	CGAGG	CGCCCT	CCCA	CTTACT	ACTG	CTG	CACTGTGG	AGG	1950					
DB	1964	--GAG	GTGTGA	CCCTG	ATG	AG	ACCC	CTGA	CA	CCCTG	CTACTG	CCAGCTGG	AGC	2021				
QY	1951	CCAGT	GCCTG	CTAG	CTTCA	CCG	AG	CAGCTG	GG	CCGCTTTG	CCCTGTG	GGGAGG	CCC	2010				
DB	2022	CCAGG	CCCTGT	CA	CTCTG	CTG	CA	CTG	GG	CACTAC	GTGTT	TCAG	GGGCGAGT	CC	2081			
QY	2011	TCAGC	GTGG	CTGG	CCCA	AGG	CCCTCA	AGTCTT	CTGTTT	GG	CGCGTGG	CTGC	CACT	2070				
DB	2082	ATTCC	CGCTCAG	AGTCA	AGCGG	CTCC	AGTGG	CCG	CTTTC	GG	CCCGCC	CTG	CACT	2141				
QY	2071	CCCTCG	AGTACA	CACTCC	GGTCTACT	CGCTGC	ATG	ACAC	CCAC	AGATG	CACTCA	AGGAGG	2130					
DB	2142	CCCTG	AGTAC	AGCTTCC	GGG	GTCTG	CTG	CTG	AG	GACAC	CGCTTAG	CACTGA	AGGAGG	2201				
QY	2131	TGTTG	CAGCTT	GGAG	AGCAGCT	GGG	GGG	CAGCTG	ATCC	AGG	AGCC	ACGGG	CTCG	CACT	2190			
DB	2202	TGCTG	AGCTGG	AGCG	GA	CTCTGG	CGG	ATCTT	GGT	GG	AGG	CGG	AA	CGCT	ATGT	2261		
QY	2191	TCAGG	ACAGT	TACCA	CACTCG	CGCTAT	CCAT	CA	AGATGT	CGCC	AGCTCC	CTGTG	GA	2250				
DB	2262	TCAGG	ACAGT	TACCA	CACTCG	CGCTAT	CCAT	CA	AGATGT	CGCC	AGCTCC	CTGTG	GA	2321				
QY	2251	AGAGT	AAGCT	CTCTG	TCAGCT	ACCA	GGAGAT	CC	CTTTT	AT	CA	CACTT	GG	AA	TCGC	AGC	2310	
DB	2322	GGAG	CAAG	CTGTG	CCCA	ATAC	AGG	AT	CCCC	CTT	AT	CA	ATTT	GG	AGTGG	CA	2381	
QY	2311	AGCGT	ATCTT	GCACTG	CACTT	CA	CCCTGG	AGG	GTGT	CAG	CCCA	CACTAG	TG	CACTGG	2370			
DB	2382	AGAA	G	CCCTCC	ACTTTC	CACTGG	AGG	GCACAG	CTT	GGC	CTTCC	ACAG	GTCA	2441				
QY	2371	CTGTCA	AGCTGTGG	GTGG	AGTGTG	AGG	CGG	CGAG	CTT	CA	CACTT	GG	AA	TCGA	2430			
DB	2442	CTTGA	AGATCTG	CGTGG	CGG	CAAGT	GG	AA	GGG	AGG	GCAG	ATAT	TC	CAG	GTGC	ATACCA	2501	
QY	2431	ACAT	CACCA	AG--	--GAC	ACA	AG	GT	TTG	CTG	AGCTG	CTGG	CTCTG	AG	AGTGA	AGCGGGG	2487	
DB	2502	CTCTG	CAG	AGAC	CGTGT	GCCT	CCCTG	GA	CACTCT	CTGT	CTG	CTG	CTG	CTG	CTG	CTG	CTG	2561
QY	2488	TTCC	AG	CCCTGTG	GG	CCCC	CA	GTCTT	CA	AGAT	CCCC	CT						

Search completed: October 6, 2004, 18:16:15
Job time : 1359 secs

Db	2622	GCAACAGCCTAGAGTGGCCCACTCAGCGGCGCAATGACTGGCGATGTTAGCACAGAAGC	2681
Qy	2608	TCCACCTGGACAGCCATCTCAGCTTCTTTTGCCTCCAGACCCAGGCCCAAGCCATGATCC	2667
Db	2682	TCCTATGGACCGGTACCTGAAATCTTTTGCCACCAAGCGAGCCCAACGGGTGTGATCC	2741
Qy	2668	TCAACCTGTGGGAGCGCGGCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAG	2727
Db	2742	TGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGACCTCAACAGCCTGGCGAGTGCCT	2801
Qy	2728	TGCTGGACTGGGCCAGCCAGACGCTTTCACAGTGTGGAGGCTGAGTGTGA	2786
Db	2802	TGAGAGAGATGGGCAAGAGTAGAGATCTGGTGCCTGTGCCCAACGACGGGAGCTGTGA	2860

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2004, 11:46:57 ; Search time 7179 Seconds
(without alignments)
11983.973 Million cell updates/sec

Title: US-09-970-944-1

Perfect score: 2881

Sequence: 1 agctggggtccgggtgag.....ccttccccacacgggggaga 2881

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_esttc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_ges1:*
29: gb_ges2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907.8	31.5	2802	29	AY406491 Homo sapi
2	898.6	31.2	2791	29	AY406493 Mus muscu
3	871.6	30.3	3790	11	AK031655 Mus muscu
4	818.4	28.4	3866	11	AK018177 Mus muscu

5	763.2	26.5	2532	29	AY411747	AY411747 Homo sapi
6	763.2	26.5	2802	29	AY406492	AY406492 Pan trogl
7	756.2	26.2	1034	12	BI758231	BI758231 603029876
8	726.4	25.2	2532	29	AY411749	AY411749 Mus muscu
9	726.2	25.2	909	13	BX364574	BX364574 BX364574
10	722.2	25.1	788	9	AI951556	AI951556 wv36f04.x
11	711.2	24.7	1532	11	BC033727	BC033727 Homo sapi
12	710	24.6	818	12	BI818609	BI818609 603033362
13	709.8	24.6	843	13	BX348193	BX348193 BX348193
14	700.2	24.3	939	13	BX345406	BX345406 BX345406
15	679.4	23.6	1040	13	BX345407	BX345407 BX345407
16	664.2	23.1	1201	13	BX422753	BX422753 BX422753
17	664	23.0	889	13	BQ691915	BQ691915 AGENCOURT
18	653.6	22.7	934	10	BF311804	BF311804 601897316
19	648.6	22.5	2532	29	AY411748	AY411748 Pan trogl
20	646.4	22.4	859	10	BF311896	BF311896 601897733
21	644.2	22.4	874	13	BQ689148	BQ689148 AGENCOURT
22	604.4	21.0	756	13	BU612387	BU612387 UI-M-EWO-
23	596.4	20.7	662	9	AL516580	AL516580 AL516580
24	579	20.1	2775	29	AY401471	AY401471 Mus muscu
25	573.4	19.9	604	10	BE314370	BE314370 601147261
26	573.4	19.9	2775	29	AY401469	AY401469 Homo sapi
27	554.6	19.3	1175	10	BF530640	BF530640 602071931
28	548	19.0	788	14	CA317532	CA317532 UI-M-FW0-
29	539.4	18.7	678	14	CA749784	CA749784 UI-M-FW0-
30	524.4	18.2	824	12	BI737024	BI737024 AGENCOURT
31	493	17.1	1030	12	BM479478	BM479478 AGENCOURT
32	471.2	16.4	540	12	BI467799	BI467799 389583 MA
33	467.8	16.2	917	13	BX453456	BX453456 BX453456
34	463.6	16.1	1028	13	BX331575	BX331575 BX331575
35	449.8	15.6	987	13	BX452510	BX452510 BX452510
36	445	15.4	2507	29	AY401470	AY401470 Pan trogl
37	433	15.0	471	13	BX282095	BX282095 BX282095
38	430.6	14.9	675	14	CA315487	CA315487 UI-M-FW0-
39	424.4	14.7	843	12	BG913440	BG913440 602811321
40	424	14.7	780	12	BG470852	BG470852 602511705
41	416	14.4	718	14	CA318180	CA318180 UI-M-FW0-
42	415.4	14.4	1147	13	BU840446	BU840446 AGENCOURT
43	413	14.3	960	12	BG329812	BG329812 602429458
44	411.6	14.3	1111	12	BG298307	BG298307 602397080
45	411.4	14.3	749	14	CF735417	CF735417 UI-M-HB0-

ALIGNMENTS

RESULT 1	AY406491	2802 bp	DNA	linear	GSS 15-DEC-2003
LOCUS	AY406491	Homo sapiens UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence.			
DEFINITION	AY406491	Genomic survey sequence.			
ACCESSION	AY406491.1	GI:39762465			
VERSION	AY406491.1	GI:39762465			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1	(bases 1 to 2802)			
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Inferred nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2	(bases 1 to 2802)			
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Direct Submission				

JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA									
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.									
FEATURES	Location/Qualifiers									
source	1..2802									
	/organism="Homo sapiens"									
	/mol_type="genomic DNA"									
gene	/db_xref="taxon:9606"									
	<1..>2802									
	/gene="UNC5C"									
	/locus_tag="HCM2575"									
ORIGIN										
Query Match	31.5%; Score 907.8; DB 29; Length 2802;									
Best Local Similarity	60.3%; Pred. No. 3.9e-142;									
Matches 1660; Conservative	0; Mismatches 1014; Indels 78; Gaps 7;									
Qy	98	GC	CGG	CGCTGT	GC	CAG	CGCTCT	CGGCATAG	CTCTCGCGCTTGGCTCGCGGCTCGGG	157
Db	66	GCT	CGTCTACT	CGCCCTTGGCCCTGCTCAGCGCAGCGGCACTGGCTCCGCCGCCAAGA	125					
Qy	158	TG	CCACAGAGTCC	ACCGTGGCCAAACCCAGTGCCTGTGGTCCAAACCCGACCTGTCTTC	217					
Db	126	TGATGACTTTT	TCATGAATCC	CAGAAACTTTTCCCTTCATCCACCTGAGCCTCTGCC	185					
Qy	218	CCA	CTTCTGTGTGAGCCG	CAGGATGTGTACATCGTCAAGAAACGACGAGTGTGCTGTGT	277					
Db	186	ACA	TTTCTTATTGAGCCT	GAAGAAGCTTATTTGTGAAGAAATAGGCTGTGAACCTGTA	245					
Qy	278	GTG	CAAGGCCGTG	CCGCCACG	CAGATCTCTTCAAGTGCACCGGAGTGGGCGCCA	337				
Db	246	CTG	TAAAGCAAGCCCTG	CCGCCCAAGATCTATTTCAGTGTAAATAGTGAATGGGTTCA	305					
Qy	338	GGT	GGACCACTGATCG	AGCCAGCACAGACGGAGACGTGCTGAGCGCACCATGAGAGGT	397					
Db	306	GAAG	ACCACACTAGTAGATGA	AAAGTAGTAACACTTCCGGTCTCATTTGTCGGGAAGT	365					
Qy	398	CCG	CATTAAATGTC	TCAAGGCAGCAGTTCGAGAAGGTGTTCCGGGTGGAGGAATACTGGTG	457					
Db	366	GAG	CATTGAGATTTCG	CCACAGCAAGTGGAAACTCTTTGGACCTGAAGATTACTGGTG	425					
Qy	458	CA	GTGCGTGGATG	GAAGTCTCTCGGCACACCAAGAGTCAAGAGGCTCATCCGCAT	517					
Db	426	CCAG	TGTGGCCCTGGAGCTCG	CGGGTACCAAAAGAGCGGAAGCGGTATGTGCGCAT	485					
Qy	518	AGC	-----	CAGATTGGC	CAAGAACTTCGAGCAGGAGCGGTGGCCAAAGGAGGTGCCCT	571				
Db	486	TGCA	TNNNNNNNN	NCGGAAGACATTTGAGCAGGAA	CCCTTAGAAGGAAGTGTCTTT	545				
Qy	572	GGAG	CAGGCATCGTGCT	CGCTCGCTCCACCGGAGGGCATCTCTCCAGCGGAGGTGGA	631					
Db	546	GGAA	CAGGAAGTCTTACT	CCAGTGTGACCACTGAAGGGATCCCAAGTGGGTGAGGTGGA	605					
Qy	632	GTG	GCTCCGGAA	CAGGAGCTCGTGGA	CCCGTCCCTGGACCCCATGTATATCATCA	691				
Db	606	ATGG	TTGAAAAATGA	GAACAATAATTGATCCCGTTGAAGATCGGAATTTTATTACTAT	665					
Qy	692	GGAG	CACAGCTGGTGGT	CGGACAGSCCGCTTGTGACAGCGCAACTACACCTCGGT	751					
Db	666	TGAT	CACAACTCAT	TAAGCAGGCCCGATCTCTGATATGCAATTAACACCTGTGT	725					
Qy	752	GGCC	AGAAACATCGTGG	CACGTGCCGCGAGCCCTCCGCTGTGTCATCGTCTACGTGAA	811					
Db	726	TGCC	AAAAACATTGTTG	CCAAAGAGGAAAAAGTACAACTGCACTATGTCTCA	785					
Qy	812	CGT	GGTGGTTCGAC	CTGTGAC	CGAGTGGTTCGCTCTGCAGCGCACGTGTGGCGCGGCTG	871				
Db	786	CGGT	GGCTGCTCCAC	CTGGA	CGAGTGGTCTGTGTGTAAACAGCCGCTGTGGACGAGGTA	845				
Qy	872	GC	AGAAACGAGCCG	GAGTGCACCAACCCGGCGCTCTCAACCGGGGGCGCTTCTGTGA	931					
Db	846	TC	AGAAACGTA	CAAGGACTTGTACCAACCGGACCACTCATATGGGGGTGCCTTCTGTGA	905					

```
QY 1955 TGCTGTCTACGCTTTCACGAGAGCTGGCGCTTTTCCCTGTGTGGAGAGCGCCTCAG 2014
Db 1971 GGCCTGCGACATCTCTCAGAGAACTCAGCACCTAGCCCTGGTAGGACATTCACAC 2030
QY 2015 CGTGTCTCCGCGAGCGCTCAAGCTCTTCTGTGTGGCGGTGGCTGACCTCCCT 2074
Db 2031 CAAGCGGCTCGAAGCGCCCTCAAGCTGGCCATCTTTGGGCGCTGTCTCTCTCGCT 2090
QY 2075 CGAGTACAACTCCGGGTCTACTGCTCATGACACCCAGCATGCACTCAAGAGGCTGT 2134
Db 2091 GGAGTACAGATCCGAGTCTACTGCTGATGACACCCAGCATGCCCTGAAGAAATTTT 2150
QY 2135 GCAGCTGGAGAGAGCTGGGGGAGAGCTGATCCAGAGAGCCAGGGTCTTGCACTTCAA 2194
Db 2151 ACATCTTTGAGAGACAGATGGGAGGACAGCTCCTAGAAAGACCTAAGGCTCTCTATTAA 2210
QY 2195 GGACAGTTTACCACACCTGGCCCTTATCCATCCAGATGTGCCAGCTCCCTGTGGAAGAG 2254
Db 2211 AGGCAGACCCACAACTGGCGCTGTCAATTCAGCATATGCCCATTTCCCTCTGGAGAG 2270
QY 2255 TAAGCTCTTGTCTAGCTTACCAGGAGATCCCTTTTATCAGATCTGGAATGCGAGCG 2314
Db 2271 CAAATTCCTGCTAAATATCAGGAATTCATTTTACATGTTTGGAGTGAATCTCAAG 2330
QY 2315 GTACTTGCACTGCACCTTCACTCCCTGGAGCGTGTCCAGCCAGCACTAGTACCTGGCGTG 2374
Db 2331 AAACCTGTGCACTGCACCTTCACTCTGGAAGATTTAGCTTGAACACAGCTGGCTGTTG 2390
QY 2375 CAAGCTGTGGTGTGGCAGCTGGAGGCGGAGGCGAGCTTCCAGTCAACTTCAACAT 2434
Db 2391 CAAACTCTGTGTGGCGAGGTGGAAGAGAGAGGCGAGATCTCCAGCTCAACTGCACCGT 2450
QY 2435 CACCAAGGACACAAGGTTTCTGAGCTGTCTGGAGCTGTGGAGAGTGAAGCGGGGTCCAGC 2494
Db 2451 GTCAGAGGAACCTACTTGGCATCATTTGCGCTGTCTGGATCTCCGGAACACCATCACAC 2510
QY 2495 CTGTGTGGGCGCCAGTGTCTTCAAGATCCCTTTCCTCAITTCGGCAGAGATAATTTCCAG 2554
Db 2511 GGTACGCGGGCGCCAGTGTCTTTCAGCATCCCTCTCCCTATCCGCGAGAGCTCTGTAGCAG 2570
QY 2555 CTGTGACCCACCTGTAGCGGGGTGCGACTGGCGGACTGTGGCCAGAAATCCACCT 2614
Db 2571 CTGTGATGCCCCCAGACAGAGGCGCATGACTGGAGGATGCTGGCCCATTAAGCTGAACCT 2630
QY 2615 GGACAGCATCTCAGCTTCTTTCCTTCCAGCCAGCCAGCCAGCCAGCATGATCTCAACCT 2674
Db 2631 GGACAGTACTTGAATTACTTTGCCACCAATCCAGCCCAACTGGCGTAACTCTGATCT 2690
QY 2675 GTGGAGGCGGCGCACTTCCCAACCGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGG 2734
Db 2691 TTGGGAAGCACAGAACTTCCAGATGGAACCTGAGCATGCTGGCAGCTGTCTTGAAGA 2750
QY 2735 ACTGGGCCAGCAGAGCTGGCTCTTTCAGTGTTCGAGGCTGAGTGTGA 2786
Db 2751 AATGGGAAGACATGAACCGTGTGTCTTTCAGCAGAGAGGCGAGTATTA 2802

RESULT 2
AY406493
LOCUS
DEFINITION Mus musculus UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence,
AY406493
ACCESSION AY406493
VERSION AY406493.1 GI:39762467
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2791)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
```

```
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 2791)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1. 2791
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2791
/gene="UNC5C"
/locus_tag="HMC2575"
ORIGIN
Query Match 31.2%; Score 898.6; DB 29; Length 2791;
Best Local Similarity 61.5%; Pred. No. 1.4e-140;
Matches 1635; Conservative 0; Mismatches 939; Indels 83; Gaps 9;
QY 190 TGCTGTGTGCCAACCGGACCTGCTTCCCACTTCTCTGTGTGAAGCGCGTGTGGAGCCGAGGATGTGTACA 249
Db 158 TTCCATCTGACCCACCTGAGCCATTGCCACATTCCTCATTTAGCCCGGAGAGCTTACA 217
QY 250 TCGTCAAGAACACGACGATGCTCTTGTGTGAAGCGCGTGTGGAGCCGACCATCTTCT 309
Db 218 TTGTGAAGAACACGCTGTGAACCTGTATTTGAAGCGAGCCCTGCACCATCTACT 277
QY 310 TCAAGTGCAACGGGAGTGGGTGCGCCAGTGGACCAACGTCGATCGAGCGCAGCACAGACG 369
Db 278 TCAAGTGCAACGCGAGTGGGTTCATCAGAAAGCACCACTAGTAGAGAGTAGATG 337
QY 370 GGAGCAGTGGTGAAGCGCACCATGAGGCTCGGCTTAATGTCTCAAGCGAGCAGGTGAGA 429
Db 338 AAACCTCTGTCTAATTTGAGAGAGTGAAGCATTCAGATTTTCACGCCAGCAGGTGGAGG 397
QY 430 AGGTGTTCCGGCTGGAGGAATACTGTGCGCAGTGTGCGATGCGAGCTCTCGGGCACCA 489
Db 398 AACTGTTTGGGCTGAAGATTACTGTGCGCAGTGTGTGCGCTGAGAGCTCAGAGGCACTA 457
QY 490 CCAAGAGTCAGAAAGGCTATCCCGCATAGCCAG-----ATTGCGCAAGAACTTCGAGC 543
Db 458 CGAAGAGTCGGAAGGCATACGTGCGCATTTCCGATCTGCGGAAGACATTCGAGC 517
QY 544 AGGAGCGCTGCGCAAGGAGTGTCCCTGGAGCAGGCGCATGTGTGCTGCCCTGCGCTCCAC 603
Db 518 AGGAACCTTTGGGAAGGAAGTGTCTTGGAGCAGGAAGTCTTACTCCAGTGTGCGCCAC 577
QY 604 CGGAGGCGATCCCTCCAGCGGAGTGGAGTGGCTCGGAAACGAGGACCTGGTGGACCCCT 663
Db 578 CTGAAGGAGTACC-----GTAGATAGCTAAAGATGAAGACATAATTGATCTCG 626
QY 664 CCTCGAAGCCCAATGTATACATCAGCGGGAGCACAGCCTGGTGTGGACAGGCGCGCC 723
Db 627 CTGAAGATCGGAACCTTTTATATTACTATCGATCACAACCTGATCATCAAGCAAGCCCGAC 686
QY 724 TTGCTGACACGCGCACTACACCTCGTGGCCAGAAACATCTGTGGCAGCTGCCGAGCG 783
Db 687 TCTCAGATACAGCAAAATATACCTGTGTGCAAAAATATTGTTGCCAAGAGAAAAAGCA 746
QY 784 CCTCGCTGTGTCTCATCGTCTACGTGAACGCTGGGTGTGTCGACGTGGACCGAGTGTCCG 843
Db 747 CCACAGCCACTGTCTATCTGTATGTTAATGTGTGCTGTCCACCTGGACAGAGTGTCTG 806
```

QY 844 TCTGAGCCGACGCTGTGGCGGGCTGGCAGAAACGAGCCGGAGCTGCACCAACCCG 903
 Db 807 TGTGTAAACAGCCGCTGTGGCGGAGGATATCAGAAACGACAAAGACCTGCACCAACCCAG 866
 QY 904 CGCCTCTCAACGGGGCGCTTCTGTGAGGGGAGAAATGCCATGACCCGACCCGCTCTCT 963
 Db 867 CCCCACTCAATGTGGGGCTTCTGTGAGGGGAGAGTGTGCAGAA---AATAGCATGCA 923
 QY 964 CTCTGCTGTCTCTGTGAGCGGAGCTGGAGCCCGTGGAGCAAGTGGTGGCGCTGTGGGC 1023
 Db 924 CTACGTTATGCTCCAGTGGATGTGTGAGGGGAGAGTGTGCAGAA---AATAGCATGCA 923
 QY 1024 TGAAGTCAACCACTGGCGAGCGGTGAGTGTCTGACCCAGCAGCCCGCAACGGAGGGG 1083
 Db 984 CTGAATGCAACCACTGGCGAGGAGGAGTGTACAGCAGCCAGCCCGCAACGGAGGGTA 1043
 QY 1084 AGAGTGCACAGGCACTGACCTGACACCCGCAACTGTACAGTACCTCTGTGTACACA 1143
 Db 1044 AGACTGTGATGGCTGTCTCTCAATCCAGAACTGCACTGATGGCTGTGCATGCAGG 1103
 QY 1144 GTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGG---CCTCATCGCGTGGCCGTCT 1200
 Db 1104 CTGCTCTGACTCAGATGATGTGGCTCTCTACGTGGGGATTGTGATCGCTTAACAGTCT 1163
 QY 1201 GCCTGTCTGCTGCTGTGTCTCATCTCTCTGTTATTGCCGGAAGAGGAGGGCTGG 1260
 Db 1164 GTCTGGCGATCACTGTGTGTGGTGGCCCTGTGTGTGTATCGGAAGAACCCGTAAGTGT 1223
 QY 1261 ACTCAGATGTGGTGTACTGTGCTTCTCACTCAGCTTCCAGCCCGTGCAGATCAAGC 1320
 Db 1224 AGTCTGATCATATGATCTCTCAGCACTCAATGGCGCTTTCAGCTGTGAATCAAGG 1283
 QY 1321 CCAGCAAGCAGACAAACCCCATCTGTCTACATTCAGCCGAGGACCTCAGCACCACCA 1380
 Db 1284 CTGCAGACAAAGATCTCTGTGCTGTCCCGCTGACC-----TCACCTCAGCTGCAGCA 1337
 QY 1381 CCTACAGGAGCTCTGTCTCCCGGAGGATG-----GGCCAGCCCGAAGTTCAGCT 1435
 Db 1338 TGTACAGGGACCTGTCTATGCTGTGATGATGTCTCAGCAAAATCCCATGACCAACT 1397
 QY 1436 CACCAATGGGCACTGTCTACCCCTGGTGGGGCGCCGACACACATCTGCACCAAGCTC 1495
 Db 1398 CTCATATCTGGACCCACTACCACTTGAATCAAGTGTACACAGCTCAGTGTGTG 1457
 QY 1496 TCCACCTC-----TGAGGCGAGAGTGTGTCTCCCGCTCTCCACCC----- 1539
 Db 1458 TCACTCTCAGGATGACCTTGCAGGATCTCTCATCCAACTGTCAACCCAGATGACCCAGT 1517
 QY 1540 -----AGAACTACTTCCGCTCCCTGCCCCGAGGCA 1569
 Db 1518 CTTTGTAGAGAAATGAGGCCCTTAACTGAAAGAACAGAGCTCTGCAAGACAGACTGACC 1577
 QY 1570 CCAGCAACATGACCTATGGGACCTTCACTTCTCTGGGGCGGGGTGATGATCCCTAATA 1629
 Db 1578 CATCTGCACAGCATTTGTAACCTTCACTCTCTTGGGGGTCACTCATCATCTCTAATT 1637
 QY 1630 CAGTATCAGCTCTCTATCCCCGAGATGCAATACCCGAGGAGAGATCTATGATGATCT 1689
 Db 1638 CAGGAGTAAGCTTGTGATTCCTCGGTGGGGCAATCTCTCAGGGGAGAGCTCTATGAAATGT 1697
 QY 1690 ACCTCAGCTGCAGAGCCGAGAGCTGAGTGTCCCTAGCTGTGCTGAGACCTGTC 1749
 Db 1698 ATGTGACTGTACACAGGAAGAAATATGAGGCCCTCCATGGAAGACTCTCAGACCTTAC 1757
 QY 1750 TGAGTCCCATCGTTAGCTGTGAGCCCTCGCTGCTCTCAACCCGCGCAGTCAATCTGTGG 1809
 Db 1758 TTACCCCTGTGTGAGCTGTGGCTCTCTGAGGCTCTCTGACCCGCGCTGTCTATCTCA 1817
 QY 1810 CTATGAGACACTGTGGGAGCCAGCCCTGACAGCTGGAGCTGGCTGCTCAAAAGAGT 1869
 Db 1818 CTCTGATCACTGTGAGACCCAGCAGCCAGGAGTGTGAAGATCAGCTCAAAACACAGG 1877
 QY 1870 CGTGGAGGGGAGCTGGGAGCAGGATGTCTGACCTGTGGCGGAGGAGCGCCCTCCACCC 1929

Db 1878 CAGTGACGGGACAAATGGGAGGATGTGTGGTG---CTTGGGAGGAGAACTTTCACAAACC 1934
 QY 1930 TCTACTATGTCAGCTGGAGGCGAGTGTCTGCTGTCTTCCAGGAGAGCTGGGCGGCT 1989
 Db 1935 CTTGTTTACATTCAGCTGGATGACAGGGCTTGGCATATCTCTCAGAGAACTCTCAGTACCT 1994
 QY 1990 TTGCCCTGTGGGAGAGGCCCTCAGCTGTGGCTGCCGCAAGCCCTCAAGCTCTCTGT 2049
 Db 1995 ATGCCCTGTGGGAGTCCACCAAGCAGCTGCCAAGCCTCTTAAACTGSCCATCT 2054
 QY 2050 TTGGCGGGTGGGCTCCACCTCCCTCGAGTACAAATCCGGGTCTACTGCTGCTGATGACA 2109
 Db 2055 TTGGGCCCCCTGCTCTCTTCCCTGGAGTACAGCATTAGAGTCTACTGCTGGATGACA 2114
 QY 2110 CCACATGCACTCAAGGAGTGGTGCAGTGGAGAGCAGCTGGGGGACAGCTGATCC 2169
 Db 2115 CACAGATGCCCTGAAAGAGTCTTCAACTGGAGAGGCAATGGGAGGACAGCTCTAG 2174
 QY 2170 AGGAGCAGCGGCTCTGCACTTCAAGGACAGATTACCAACACCTGCGCTATCCATCCACG 2229
 Db 2175 AAGAACCCAGGCTCTTCAATTTAAGGAGCAGATCCACACCTGCGCTCTCTATTATG 2234
 QY 2230 ATGTGCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTACGTACACAGGAGATCCCTTTT 2289
 Db 2235 ACATGCCCATTCCTCTGGAAGAGCAAAATGTCTGCTAAGTATCAGAAATTCATTTT 2294
 QY 2290 ATCAGATCTGAATGACGAGCGGTACTTGGCACTGCACCTTCAACCTCGAGGCTGTCA 2349
 Db 2295 ACCACATCTGAGTGGCTCTCAAGAAACCTCCACTGCACCTTCACTCTGGAAGACTCA 2354
 QY 2350 GCCCAGCACTAGTACCTGCGCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGGACGCGC 2409
 Db 2355 GCCTAACAACAGTGGAACTGTTTGCAAACTCTGTGTGGCGAGGTTGGAAGAGAGGGC 2414
 QY 2410 AGAGCTTTCAGCATCAACTTCAACATCAACAGGACACAAGTGTGTGAGCTGTGGCTC 2469
 Db 2415 AGATCTTCCAGCTCAACTGTACTGTGTGAGAGGAACTTCTGGCATCGACTTACCTCTCC 2474
 QY 2470 TGGAGATGAGCGGGGCTCCAGCCCTGTGGGCGCCAGTGCCTTCAAGATCCCTTCC 2529
 Db 2475 TGGACCTGTGTAGTACCATCACCCTGTCAACCGGACCAAGTGTTCAGCATCTCTCC 2534
 QY 2530 TCATTGGCAGAGATTAATTTCCAGCTGGAACCCACCTGTAGGGGGGTGCGAGCTGGC 2589
 Db 2535 CTATCCGCGAGAGCTATGAGAGCTGTGATGCCCTCAAAAGAGGCGCATGACTGGA 2594
 QY 2590 GGACTCTGGCCGAGAAACTCCACTGGAAGCCATCTCAGCTTCTTTGCTTCCAGGCCA 2649
 Db 2595 GGATGCTGGCCCAATAACTCAACCTGGACAGTACTTGAATTAATTTGGCCCAAAATCGA 2654
 QY 2650 GCGCCACAGCATGATCTCAACCTGTGGGAGGCGCGCACTTCCCAACGGCAACCTCA 2709
 Db 2655 GCGCAACTGGCGTAACTCTGGATCTTTGGGAAGCAGAGACTTCCAGATGGAAACCTGA 2714
 QY 2710 GCCAGCTGGCTGACAGCTGGCTGGAGTGGCCAGCCAGAGCTGGGCTCTTCAAGTGT 2769
 Db 2715 GCATGCTGGCAGCGCTCTTGGAAAGATGGAAGACATGAGACAGTGTGTCTTGGCAG 2774
 QY 2770 CGGAGGCTGAGTGTGA 2786
 Db 2775 CAGAAGGACAGTATTGA 2791

RESULT 3
 AK031655
 LOCUS
 DEFINITION
 Mus musculus 13 days embryo male testis cdna, RIKEN full-length
 enriched library, clone:6030473H24 product:unc5 homolog (C.
 elegans) 3, full insert sequence.
 AK031655
 VERSION
 AK031655.1 GI:26327502
 KEYWORDS
 HTC; CAP trapper.


```

SOURCE
ORGANISM Mus musculus (house mouse)
Bukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
2
REFERENCE Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
AUTHORS Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
3
REFERENCE Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
AUTHORS Kono, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, I., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
4
REFERENCE The RIKEN Genome Exploration Research Group Phase II Team and the
AUTHORS FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
PUBMED 11076861
5
REFERENCE The FANTOM Consortium and the RIKEN Genome Exploration Research
AUTHORS Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 12401370
6
REFERENCE Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
AUTHORS Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, K., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, I., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
FEATURES Location/Qualifiers
source 1..3790

```

CDS

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM,DB:6030473H24"
/db_xref="MGI:2394911"
/db_xref="taxon:10090"
/clone="6030473H24"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
118..2970
/note="unnamed protein product; putative
unc5 homolog (C. elegans) 3 (MGD|MGI:1095412,
GB|NM_009472, evidence: BLASTN, 99%, match=464)"
/codon_start=1
/protein_id="BAC27495.1"
/db_xref="GI:26327503"
/translation="MKRLRATAARCLGIGYLLQMLVLPALLSASGTGSAQDDDE
FFHELPDPSPDPPEPLPHLEPERAYIVKNPKVNLKASPAATQIYFKCNSEWHQ
KDHVDERVDTSGLIVREVISIISRQVEELFEPEDYWCQCVASAGATYTKRKAIV
RIAYLRKTEQEPGLKEVLEQVLLQCRPEGIPIVAEVEWLNEDIIDPAERNFYI
TIDHNLIIKOARLSDTANTCVAKNIVAKRSTATVIVYVNGSNWTWTSVNCNSC
RGVQKRTCTNPAPLNGAFCEGVSQVKIACITLCPDVGRWISWKSMTCTGCTETH
WRRECTAPAKNGKDCDLVLQSKNCTDGLCMQGFIPDIETHEHPQNEYGFSSAPD
SDVALYGVIVATVCLAITVVVALFVYRKNRDFESDIIDSSALNGQPQVNIKAA
RODLAVPDLTSAAMYGPVVALHDVSDKIPTMNSPILDPNLKIKIVYNSGAVT
PODLAEFSKLSPOQSLLENALNKNOSLAROTDPSCTAFGTENSIJGHLIIPN
SGVSLIPAGAIPOGRVYEMVTVHRKENRPMEDSOIILLTPVSCGPGGALLRPV
ILTLHCADPSTEDWKIOLKNQAVQGWEDVVVVEENFTTCYIOLDAEACHILEN
LSTYALVGGSTKAAAKRLAIFGLPCSSLEYSIRVYLCCTDQALKEVLEHTEM
GGOLLEPRLAKRFKSIHNLRSIHDIASLWKSLLAKYQETPFPHIWSGQSNLHC
TPTLESLNTVELVKLCVROVEGGOIFQLNCTVSEPTGIDLPDLPASTITVT
GSAFSIPLPIROKLCSLSDAPOTRGHWRMLAHKLNLDRLYNLFATKSPSTGVILDL
WEAQNPDPDGNLSMLAAVLEEMGRHETVVIABEGQY"

```

ORIGIN

```

Query Match 30.3%; Score 871.6; DB 11; Length 3790;
Best Local Similarity 60.6%; Pred. No. 5-se-136;
Matches 1641; Conservative 0; Mismatches 944; Indels 123; Gaps 8;
QY 190 TGCTGCTGCCAACCGGACCTGCTTCCCACATCTCTGGTGGAGCCCGAGGATGTGTACA 249
Db 275 TTCCATCTGACCCACCTGAGCCATTCGCCACATCTCTCATTTAGCCCGGAGGAGCTTACA 334
QY 250 TCGTCAAGAACAAAGCCAGTGTCTTGTGTCGAAGCCCTGCCCGCCAGCGATCTTCT 309
Db 335 TTGTGAAGAACAAAGCTGTGAACCTGTATTGTAAGCCAGCCCTGCCACCCAGATCTACT 394
QY 310 TCAAGTGCACCGGGAGTGGGTGGCCAGGTGGACCACTGATCGAGCGCAGCAGACAG 369
Db 395 TCAAGTGCACACGAGTGGGTGTTTCATCAGAAGGACCACTAGTAGCAGAGAGTAGATG 454
QY 370 GGAGCAGTGTGTGAGCGGACCATGGAGTCCCGCATTAATGTCTCAAGGACGAGGTGGAGA 429
Db 455 AAACCTCTGGTCTAATTTGTGAGAAGGTGAGCATTTGAGATTTTCAGCCGACGAGTGGAG 514
QY 430 AGGTGTCCGGTGGAGGAATACTCGTCCAGTGTGGTGGCAGTGGAGTCTCTCGGACCA 489
Db 515 AACTGTCTGGGCTGAAGATTACTGTGTCAGTGTGTGGCTGGAGCTTACGACGAGCACTA 574
QY 490 CCAAGAGTCAGAAGCCCTACATCCGATAGCCAGATTCCGCAAGAACTTCGAGCAGGAGC 549
Db 575 CGAAGAGTCGAAGGCATACGTCGCGATTCGCGATATCTCGGAAGACATTCGAGCAGGAG 634
QY 550 CGCTGGCCAAAGGAGTGTCCCTGGAGCAGGCGATCTGCTGCCCTGCGCTCCACCGGAG 609
Db 635 CTTGGGAAAGGAAGTGTCTTGGAGCAGGAGGATCTTACTCAGTGTGCGCCACCTGAAG 694
QY 610 GCATCCCTCCAGCCGAGGTGGAGTGGCTCCCGAAGCAGGAGCTGTGTGACCGGTCCCTGG 669
Db 695 GGATCCCGAGTGGCTGAGGTAGATGGCTAAAGAAATGAAGACATAATTGATCTCTCTGAAG 754

```

QY 670 ACCCCAAATGATATACATCAACGCGGAGCAAGAGCTGTGTGTGCGAAGGCGCGCTTGCTG 729
 Db 755 ATCGGAATTTTATATTAATCATGATCAACAACCTGATCATCAAGCAAGCCGACTCTCAG 814
 QY 730 AACGCGCAACTACACTGCTGGTGGCCCAAGAACTCTGGCACGTCCGCGGAGCGCTCCG 789
 Db 815 ATACAGCAAAATATATCTGTGTGGCCAAATAATTTGTTCGAAGAAAGAACACACAG 874
 QY 790 CTGCTCTCATCTGCTACGTGAACGCTGGGTGGTGGACCTGGAGCCGAGTGGTCCGCTGCA 849
 Db 875 CCACTGTCTATCTGTATGTTAATGGTGGTGGTCCACCTGGACAGAGTGGTCTGTGTGA 934
 QY 850 GCGCCAGCTGTGGCGCGCTGGGAGAAACGAGAGCGGAGCTGCACCAACCGCGCGCTC 909
 Db 935 ACAGCGCTGTGGCGGAGGATATCAGAAAACGCAAGAACTGCACCAACCGAGCCGCCAC 994
 QY 910 TCAACGGGCGCTTTCTGTGAGGGGAGAGTGTCCATGACCCACCGCTCTCTCTCTGC 969
 Db 995 TCAATGGTGGGGCTTCTGTGAGGGGAGAGTGTGCAGAA---AATGATGCACTACGT 1051
 QY 970 TTGTCTCTGTGGAGCGGAGCTGGAGCCGCTGGAGCAAGTGGTGGGCTGTGGGCTGACT 1029
 Db 1052 TATGTCCAGTGGATGTTAGTGGACTTTCATGGAGCAATGGTCAACCTGTGGGACTGAAT 1111
 QY 1030 GCACCCACTGGCGAGCGGTGAGTGTCTGACCCAGCACCCCGCAACGAGGGGAGGAGT 1089
 Db 1112 GCACCCACTGGCGAGGAGGAGTGTACAGCACAGCCCGCAAGAACGGGGGTAAAGACT 1171
 QY 1090 GCCAGGCACTGACTGGACACCGCAACTGTACAGTACCTCTGTGTACACAG-----TGCTTCTG 1144
 Db 1172 GTGATGGCCTGGTCTCCATCCAAATCCAAAGAACTGCACTGATGGGTGTGATGCGGGATTCA 1231
 QY 1145 -----TGCTTCTG 1152
 Db 1232 TTTACCCCAATTCAACTGAGCACAGACCCCAAGATGAATGGATTTTCTTCTCTCTG 1291
 QY 1153 GCCGTGAGGAGTGGCCCTCTATGTGG---CCTCATGCGGTGGCGCTGCTGCTGTGCC 1209
 Db 1292 ACTCAGATGATGTGGCTCTCTACGTGGGGATGTGATGCTGTAAACAGTCTGTCTGGGA 1351
 QY 1210 TGCTGCTGTGTCTCTATCTCTGTTTATTCGCGAAGAGGAGGGGTGGACTCAGATG 1269
 Db 1352 TCATGTGTGGTGGCCCTGTGTGTATCGAAGAACCCCGTGACTTTGACTGTGACA 1411
 QY 1270 TGGCTGACTGTCTTCACTTCACTCAGGCTTCCAGCGCGTGCAGATCAAGCCCGAGCAAG 1329
 Db 1412 TCAITGACTCTCAGCACTCAATGCGGCTTTCAGCCTGTGAACATCAAGGCTGCCAGAC 1471
 QY 1330 CAGACACCCCGATCTGTCCACCATCCAGCGGAGCTCAGCACCAACCACTACACAGG 1389
 Db 1472 AAGATCTCTGGCTGTCCCGCTGACC-----TCACCTCAGTGCAGCCATGTACAGGG 1525
 QY 1390 GCAGTCTCTGTCCCGGAGGATG-----GGCCAGCGCCCAAGTTTCCAGCTCACCAGTGG 1444
 Db 1526 GACCTGTCTATGCTCTGTGATGATGTCTCAGACAAATCCCAATGACCAACTCTCCAAATC 1585
 QY 1445 GCACCTGTCTAGCCCGCTGGTGGGCGCGGCACACATGACCAACCACTCTCCACCTC 1504
 Db 1586 TGGACCCCACTACCCAACTTGAATAAATAAAGTGTACACAGCTCAGGTGTCTCACTCTC 1645
 QY 1505 -----TGAGGCCGAGAGTGTCTCTCCCGCTCTCCACCC----- 1539
 Db 1646 AGGATGACTCTCCGAGTCTCATCCAACTGTCCCGAGATGACCCAGTACCAGTCTTGCTAG 1705
 QY 1540 -----AGAACTACTTCCGCTCTCCCTGCGCGAGCCACAGCAACA 1578
 Db 1706 AGAATGAGCCCTTAACCTGAAGAACAGAGCTCGCAAGACAGACTGACCCCATCTGCA 1765
 QY 1579 TGACCTATGGGACCTTCACTCTCGGGGCGCGCTGATGATCCCTAATACAGGATCA 1638
 Db 1766 CAGCAATTTGGTACCTTCAACTCTCTTGGGGGTCACTCATCATCTCTAATCAGGAGTAA 1825
 QY 1639 GCCTCTCATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCACGC 1698

Db 1826 GTTGTCTGATTCGCGTGGGCCATTTCTTCAGGGGAGAGTCTATGAAATGTATGTGACTG 1885
 QY 1699 TGCACAGCCGGAAGAGCTGAGGTTGCCCTAGCTGGCTGTGAGACCCCTGCTGAGTGCCTCA 1758
 Db 1886 TACACAGGAAGAAATATAGGCGCCCATATGAGAGACTCTCAGAGCTCTTACCCCTG 1945
 QY 1759 TGGTGTAGCTGTGAGACCCCTGCGTCTCTCACCCGCGCAGTCACTCTGGCTATGAGACC 1818
 Db 1946 TGTGTAGCTGTGGGCTCTCTGGAGCTCTGTGACCGCGCTGTCACTCTCACTCTGCATC 2005
 QY 1819 ACTGTGGGAGGCGCCAGCTGTGAGCCTGTGAGCCTGTGCGCTCAAAAAGAGCTGTGGAGG 1878
 Db 2006 ACTGTGCAGACCCAGCACCGAGGACTTGAAGATCCAGCTCAAAAACAGGAGTGGAGG 2065
 QY 1879 GCAGCTGGAGCAGAGTGTCTGCACCTGGGCGAGGAGGCGCCCTCCCACTCTACTACT 1938
 Db 2066 GACATGGGAGGATGTGTGGTG---GTTGGGGAGGAGAACTTCAAAACCCCTGTTACA 2122
 QY 1939 GCCAGCTGGAGGCGAGTGTCTGCTTTCACCGAGCAGCTGGGCGGCTTTGCGCTGG 1998
 Db 2123 TTCACTGTGATGAGAGGCTTGCATATCTTCACAGAGAACCTCAGTACCTATGCGCTGG 2182
 QY 1999 TGGGAGAGGCGCTCAGCGTGGCTGCGCGCAAGCGCTCAAGCTGTCTTCTGTTGCGCGG 2058
 Db 2183 TTGGGAGTCCACCAACCAAGCAGCTGCCAAGGCTCTTAACTGGCCATCTTTGGGCGCC 2242
 QY 2059 TGGCTGTCACTCCCTCGAGTAAACATCCGGGTCTACTGCTGTGATGACACCCAGCATG 2118
 Db 2243 TCTGTCTCTTCCCTGGAGTACAGCATTAGAGTCTACTGCTGTGATGACACAGATG 2302
 QY 2119 CACTCAAGAGGTGTGTCAGCTGAGAAAGCAGCTGGGGGACAGCTGATCAGAGAGCAC 2178
 Db 2303 CCTGAAGAGAGTCTTCAACTGGAGAGCAATGGGAGGACAGCTCCTAGNAGAACCA 2362
 QY 2179 GGGTCTGTGCACTTAAGAGCAGTTACCAACCTGCGCCTATTCATCCAGATGTGCCCA 2238
 Db 2363 AGGCTCTCTGTTTAAAGGAGCAGATCCCAACCTGCGCCTGTCTATTTCATGACATCGCC 2422
 QY 2239 GCTCCCTGTGGAAGAGTAAGCTCTTCTCAGCTACAGGAGATCCCTTTTATCACAATCT 2298
 Db 2423 ATTCCTCTGGAGAGCAATTTGCTGCTAAGTATCAGAAATTCATTTTACACATCT 2482
 QY 2299 GGAATGCGACGCGGTACTTGCATCTGCACCTTCACTGAGCGTGTGAGCGGAGTCA 2358
 Db 2483 GGAGTGTCTTCAAGAAACCTTCACTGACCTTCACTCTGAAAAGACTCAGCTTAAACA 2542
 QY 2359 CTAGTGTGCTGCGCTGCAAGCTGTGGGTGTGGAGTGGAGGCGAGCGGCGAGAGCTTCA 2418
 Db 2543 CAGTGGAACTGGTTTGAACACTCTGTGTGCGGAGGTTGAAGAGAGGCGAGATCTTCC 2602
 QY 2419 GCATCAACTTCAACATCACCAAGGACACAAAGTTTGTGAGCTGTGCTGTGAGAGTGT 2478
 Db 2603 AGCTCAACTGTACTGTGTGAGAGAACCTACTGGCATGCACTTACCTCTCTGAGACCTG 2662
 QY 2479 AAGCGGGGTGCCAGCCCTGTGGGCGCCAGTGTCTTCAAGATCCCTTCTCATTTGCGC 2538
 Db 2663 CTAGTACCATCACCACTGTCAACCGGACCAAGTGTCTTTCAGCATCTCTCTGCTATCCGC 2722
 QY 2539 AGAAGATAAATTTCCAGCTGTGACCCACCTGTAGCGGGGTGCCAGTGGCGGACTCTGG 2598
 Db 2723 AGAAGCTATGAGCAGCGCTGGATGCCCTCAAAAGAGGCGCATGACTGGAGGATGTGG 2782
 QY 2599 CCCGAAACTCCACTGAGAGCCATCTCAGTCTTCTTTCCTTCCAAAGCCAGCCCCACAG 2658
 Db 2783 CCCATAAATACTCACTGAGAGGACTTGAATTTTTCACCAAAATCGAGCCCACTG 2842
 QY 2659 CCATGATCCTCAACCTGTGGAGGCGCGACTTCCCAAGCGCAACCTCAGCCAGCTGG 2718
 Db 2843 GCGTAATCTTGATCTTTTGGGAAGCACAGAACTTCCAGATGGAAACCTGAGCATGCTGG 2902
 QY 2719 CTGACGAGTGTGCTGACTGGGCCAGCAGCTGGGCTCTTTCACAGTGTGCGAGGCTG 2778

Db 2903 CAGCCGTCCTGGAGAAATGGGAGACATGAGACAGTGTGTCTTGGCAGCAGAGAC 2962

Qy 2779 AGTGCTGA 2786
|||||

Db 2963 AGTATTGA 2970

RESULT 4
AK018177
LOCUS
DEFINITION
Mus musculus adult male medulla oblongata cDNA, RIKEN full-length
enriched library, clone:630415802 product:TRANSMEMBRANE RECEPTOR
UNC5H2 homolog [Rattus norvegicus], full insert sequence.
AK018177
VERSION
GI:12857775
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
9279253
MEDLINE
PUBMED
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
PUBMED
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Wakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, E., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
PUBMED
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (baees 1 to 3866)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, K., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toyota, Y., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GAGAGAGAGAGATCAAGAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATCTCGATTCTAATAATCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FUC 1. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

FEATURES	SOURCE
1. <u>Age</u>	1. <u>Age</u>
2. <u>Gender</u>	2. <u>Gender</u>
3. <u>Marital Status</u>	3. <u>Marital Status</u>
4. <u>Education</u>	4. <u>Education</u>
5. <u>Income</u>	5. <u>Income</u>
6. <u>Occupation</u>	6. <u>Occupation</u>
7. <u>Religion</u>	7. <u>Religion</u>
8. <u>Political Affiliation</u>	8. <u>Political Affiliation</u>
9. <u>Health Status</u>	9. <u>Health Status</u>
10. <u>Travel History</u>	10. <u>Travel History</u>
11. <u>Employment Status</u>	11. <u>Employment Status</u>
12. <u>Family Size</u>	12. <u>Family Size</u>
13. <u>Home Ownership</u>	13. <u>Home Ownership</u>
14. <u>Vehicle Ownership</u>	14. <u>Vehicle Ownership</u>
15. <u>Insurance Status</u>	15. <u>Insurance Status</u>
16. <u>Substance Use</u>	16. <u>Substance Use</u>
17. <u>Mental Health</u>	17. <u>Mental Health</u>
18. <u>Chronic Conditions</u>	18. <u>Chronic Conditions</u>
19. <u>Acute Conditions</u>	19. <u>Acute Conditions</u>
20. <u>Vaccination Status</u>	20. <u>Vaccination Status</u>
21. <u>Compliance with Medical Advice</u>	21. <u>Compliance with Medical Advice</u>
22. <u>Healthcare Utilization</u>	22. <u>Healthcare Utilization</u>
23. <u>Healthcare Costs</u>	23. <u>Healthcare Costs</u>
24. <u>Healthcare Access</u>	24. <u>Healthcare Access</u>
25. <u>Healthcare Satisfaction</u>	25. <u>Healthcare Satisfaction</u>
26. <u>Healthcare Quality</u>	26. <u>Healthcare Quality</u>
27. <u>Healthcare Equity</u>	27. <u>Healthcare Equity</u>
28. <u>Healthcare Innovation</u>	28. <u>Healthcare Innovation</u>
29. <u>Healthcare Research</u>	29. <u>Healthcare Research</u>
30. <u>Healthcare Policy</u>	30. <u>Healthcare Policy</u>
31. <u>Healthcare Regulation</u>	31. <u>Healthcare Regulation</u>
32. <u>Healthcare Ethics</u>	32. <u>Healthcare Ethics</u>
33. <u>Healthcare Law</u>	33. <u>Healthcare Law</u>
34. <u>Healthcare Economics</u>	34. <u>Healthcare Economics</u>
35. <u>Healthcare History</u>	35. <u>Healthcare History</u>
36. <u>Healthcare Future</u>	36. <u>Healthcare Future</u>
37. <u>Healthcare Trends</u>	37. <u>Healthcare Trends</u>
38. <u>Healthcare Challenges</u>	38. <u>Healthcare Challenges</u>
39. <u>Healthcare Opportunities</u>	39. <u>Healthcare Opportunities</u>
40. <u>Healthcare Solutions</u>	40. <u>Healthcare Solutions</u>
41. <u>Healthcare Best Practices</u>	41. <u>Healthcare Best Practices</u>
42. <u>Healthcare Case Studies</u>	42. <u>Healthcare Case Studies</u>
43. <u>Healthcare Research Papers</u>	43. <u>Healthcare Research Papers</u>
44. <u>Healthcare Journals</u>	44. <u>Healthcare Journals</u>
45. <u>Healthcare Conferences</u>	45. <u>Healthcare Conferences</u>
46. <u>Healthcare Workshops</u>	46. <u>Healthcare Workshops</u>
47. <u>Healthcare Seminars</u>	47. <u>Healthcare Seminars</u>
48. <u>Healthcare Webinars</u>	48. <u>Healthcare Webinars</u>
49. <u>Healthcare Podcasts</u>	49. <u>Healthcare Podcasts</u>
50. <u>Healthcare Blogs</u>	50. <u>Healthcare Blogs</u>
51. <u>Healthcare Social Media</u>	51. <u>Healthcare Social Media</u>
52. <u>Healthcare News</u>	52. <u>Healthcare News</u>
53. <u>Healthcare Analysis</u>	53. <u>Healthcare Analysis</u>
54. <u>Healthcare Commentary</u>	54. <u>Healthcare Commentary</u>
55. <u>Healthcare Debate</u>	55. <u>Healthcare Debate</u>
56. <u>Healthcare Discussion</u>	56. <u>Healthcare Discussion</u>
57. <u>Healthcare Collaboration</u>	57. <u>Healthcare Collaboration</u>
58. <u>Healthcare Partnership</u>	58. <u>Healthcare Partnership</u>
59. <u>Healthcare Alliance</u>	59. <u>Healthcare Alliance</u>
60. <u>Healthcare Network</u>	60. <u>Healthcare Network</u>
61. <u>Healthcare Community</u>	61. <u>Healthcare Community</u>
62. <u>Healthcare Association</u>	62. <u>Healthcare Association</u>
63. <u>Healthcare Organization</u>	63. <u>Healthcare Organization</u>
64. <u>Healthcare Institution</u>	64. <u>Healthcare Institution</u>
65. <u>Healthcare Facility</u>	65. <u>Healthcare Facility</u>
66. <u>Healthcare Service</u>	66. <u>Healthcare Service</u>
67. <u>Healthcare Program</u>	67. <u>Healthcare Program</u>
68. <u>Healthcare Initiative</u>	68. <u>Healthcare Initiative</u>
69. <u>Healthcare Project</u>	69. <u>Healthcare Project</u>
70. <u>Healthcare Task Force</u>	70. <u>Healthcare Task Force</u>
71. <u>Healthcare Committee</u>	71. <u>Healthcare Committee</u>
72. <u>Healthcare Board</u>	72. <u>Healthcare Board</u>
73. <u>Healthcare Council</u>	73. <u>Healthcare Council</u>
74. <u>Healthcare Commission</u>	74. <u>Healthcare Commission</u>
75. <u>Healthcare Agency</u>	75. <u>Healthcare Agency</u>
76. <u>Healthcare Department</u>	76. <u>Healthcare Department</u>
77. <u>Healthcare Division</u>	77. <u>Healthcare Division</u>
78. <u>Healthcare Unit</u>	78. <u>Healthcare Unit</u>
79. <u>Healthcare Team</u>	79. <u>Healthcare Team</u>
80. <u>Healthcare Group</u>	80. <u>Healthcare Group</u>
81. <u>Healthcare Staff</u>	81. <u>Healthcare Staff</u>
82. <u>Healthcare Personnel</u>	82. <u>Healthcare Personnel</u>
83. <u>Healthcare Professionals</u>	83. <u>Healthcare Professionals</u>
84. <u>Healthcare Experts</u>	84. <u>Healthcare Experts</u>
85. <u>Healthcare Leaders</u>	85. <u>Healthcare Leaders</u>
86. <u>Healthcare Influencers</u>	86. <u>Healthcare Influencers</u>
87. <u>Healthcare Thought Leaders</u>	87. <u>Healthcare Thought Leaders</u>
88. <u>Healthcare Visionaries</u>	88. <u>Healthcare Visionaries</u>
89. <u>Healthcare Innovators</u>	89. <u>Healthcare Innovators</u>
90. <u>Healthcare Pioneers</u>	90. <u>Healthcare Pioneers</u>
91. <u>Healthcare Trailblazers</u>	91. <u>Healthcare Trailblazers</u>
92. <u>Healthcare Game Changers</u>	92. <u>Healthcare Game Changers</u>
93. <u>Healthcare Disruptors</u>	93. <u>Healthcare Disruptors</u>
94. <u>Healthcare Challengers</u>	94. <u>Healthcare Challengers</u>
95. <u>Healthcare Reformers</u>	95. <u>Healthcare Reformers</u>
96. <u>Healthcare Activists</u>	96. <u>Healthcare Activists</u>
97. <u>Healthcare Advocates</u>	97. <u>Healthcare Advocates</u>
98. <u>Healthcare Lobbyists</u>	98. <u>Healthcare Lobbyists</u>
99. <u>Healthcare Stakeholders</u>	99. <u>Healthcare Stakeholders</</u>

```

1. 5000
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM.DB:6330415E02"
/db_xref="MGI:1910526"
/db_xref="taxon:10090"
/clone="6330415E02"
/sex="male"
/tissue_type="medulla oblongata"
/clone_lib="RIKEN full-length ex
417_3254
/ncbi="unannotated protein product;
UNC5H2 homolog [Rattus norvegicus]
FASTA, 96.5%ID, 100%length, matu
putative"
CDS

```

CDS

```

/proc/self/stat_tgid=1
/proc/self/stat_tgid=1 "
/proc/self/stat_tgid=1 "
/db_xref="GI:12857776"
/translation="MEARSQSVRSAA
BQLPYFLLEPQDAIYVKNKPVLLH
LRREVOIETVSQPOVEELFLGLEDYI
AKEVLDHVEVLLOCRBPBPVVAE
DTANYTCVAKNIYAKRRSTAATVI
APLNGGAFCEGQAFQKTAQCTVCP
GRDCSTGLLDSKNCTDGLCVLNQR
AEGVIVYRNRCDTDTIDTSSAA
IYRGPVVALQDSADKI PMTNS PLII
TYPGDFSRDTHLHLRSASLGSOA
VNGCAI POGKFVDLYLHINKASTI
CAEVIAGDWI FOLQTOAHQGHWEV
MGESYRSVAARLQIAI FAPALCTC
EPKLLFTFKDSYHNLRSLDHTPHAF
HSLASTFTCKVCRVVEGEQIYFC
KIPLSIIRQKI CSSLDAPQSDRNDWR
QDDGDLNSALAPKDSKSEMLVAF

```

ORIGIN

	Query Match	28.4%	Score 818.4	DB 11	Length 3866
	Best Local Similarity	59.6%	Pred. No. 4.9e-127		
	Matches 1623	Conservative 0	Mismatches 961	Indels 138	Gaps 9
198	GCCAACCCCGAGCTGCTCCACATCTCTGTGGAGCCGAGGAGTGTAATGCTAAG				257
540	GCCCTGGCGAGAGCTGCGGTACTTCTATTGGAGCCACGAGGACGCTACATGCTAAG				599
258	ACAAAGCCAGTGCTGTTGTGTGCAAGCCGCTGCCGCCAGCGAGATCTTCTCAAGTGC				317

Db	600	AAACAAGCCAGTGGAACTGCACCTGCAGAGCCCTCCCTGCCACGCAGATCTACTTCAAGTGT	659
QY	318	AACGGGGAGTGGGTGCGCCAGGTGGACCAACGTGATCGAGCGCAGCAAGACGGGAGCAGT	377
Db	660	AATGGCGAGTGGGTGAGCCAGAAATGACCAACGTACACAGGAGAGCCTGGATGAGGCCACA	719
QY	378	GGTAGCCGACCAATGAGAGTCCGCATTAATGTTCTCAAGGCAGCAGGTGAGAAAGTGTTC	437
Db	720	GGCTTCGGGTGCGAGAGGTGCAGATCGAGGTGTCAAGGCACAAAGTGGAGAACTCTTC	779
QY	438	GGCTCGAGGAAATACTGGTGGCCAGTGCCTGGATGGAGCTCTCGGCACCAACCAAGAGT	497
Db	780	GGCTCGAGGACTACTTGGTCAGTGGCGGTGGGCTGGAGCTCTTCGGAACTACCAAGAGT	839
QY	498	CAGAAGCCCTACATCCGCATPAGCCAGTAATTCGCAAGAACTTCGAGCAGGAGCCGTGACC	557
Db	840	CGCCGAGCCTACATCCGCATTCCTACTTTCGCAAGAACTTTGACCAAGAGCCTCTGACC	899
QY	558	AAGGAGGTGTCCTTGGAGCAGGGCATCTGTGTGCCCTTGCCTGCGTCCACCGGAGGGCATCCCT	617
Db	900	AAGAGGTACCTTTGGATCATGAGGTCTCTGTGCAGTGCGCCCAACCGGAGGAGTGGCT	959
QY	618	CCAGCCGAGTGGAGTGGCTCCGCAACGAGGACCTGGTGCACCCGTCCTCGACCCCAAT	677
Db	960	GTGGCTGAGGTGGAATGGCTCAAGATGGAAGATGTCAATTGACCCCGCTCAGAGCACTAAC	1019
QY	678	GTATACATCACGGGGAGCAGACGCTGGTGTGCGACAGGCCCGCGCTTCCTGCACGGCC	737
Db	1020	TTCTGTCTACCATTTGACCACACCTCATCATCGCCAGCGCGCGCTCTCAGACACGGCC	1079
QY	738	AATPACACTCTGCTGGCCCAAGAACATCTGTGGCAGCTGCGCGCAGCGCCTCCGCTGCTGC	797
Db	1080	AACTACACCTGTGTGGCCAAAGATATCTGTGGCCAAAGCGCGCGAGCACCGCGGCCACAGTC	1139
QY	798	ATCGTCTACGTGAACGCTGGGTGGTGCAGCTGACCGAGTGGTCCGTCGTGCAGCGCCAGC	857
Db	1140	ATCGTCTATGTGAATGAGGCTGGTCAGCTGGGCGAGGTGGTCACTCTGTTCCAAATCGC	1199
QY	858	TGTGGGCGCGCTGCGCAAAACGGAGCGCGAGTGCACCAACCGCGCGCTCTCAACGGG	917
Db	1200	TGTGGCGAGGCTGGCAGAGCGTACTCGACCTGCACCAATCCAGCCCCACTCAATGGA	1259
QY	918	GGCGCTTTCTGTAGGGGCGAATGTCCATGACCGCACCGTCTCTCTCTGCTTGTCTCT	977
Db	1260	GGCGCCTTCTGTAGGGGACAGCCCTTCCAGAGACACAGCTTGCACCAACCGGTG- --CCCA	1316
QY	978	GTGACGCGCAGCTGGAGCCGTGGAGCAAGTGTTCGSCCTGTGGGTGAGACTGCACCCAC	1037
Db	1317	GTGATGGAGCGTGGACCGAGTGGAGCAAGTGTCTGCTCTGCAGCACAGAGTGTGCGCAC	1376
QY	1038	TGGCGGAGCCTGAGTGTCTGTACCCAGCACCCCGCAAACGGAGGGGAGGAGTGCCAGGGC	1097
Db	1377	TGGCGAGCCCGAGTGCATGGCACCGCCACCCAGAAACGGAGCCGTGACTGCACGGGG	1436
QY	1098	ACTGACCTGGACACCCGCAACTGACAGTGAACCTGTGTGTACACAGTGTCTTGCGCC- --	1155
Db	1437	ACGCTACTTGACTCCAAGAACTGACCTGATGGGCTGTGCGTGTGAATCAGAGAACTCTA	1496
QY	1156	-----CTGAGGACGTGGCCCTCTATGTGGGCCTC	1184
Db	1497	AACGACCCATAAAGCACCCCTGGAGACATCGGGAGATGGGCACCTGTACGAGGCTT	1556
QY	1185	ATGCGCGTGGCGCTGTGCTGCTGTGCTTGTGTC- --CTCATCTTCGTTTATTCG	1241
Db	1557	GTGTGGCCGTCTTGTGGTGTAGCGGTCTCTATGGCCGAGGGAGTATCGTATACCGG	1616
QY	1242	CGGAAGAAAGAGGGGCTGGACTCAGATGTGGTGTGACTCGT- --CCATTCTCACTCAGGC	1298
Db	1617	AGAACTCCGGGACTTCGACACGGACATCACCGACTCTCTGTGGGCCCTCACTGGTGGC	1676
QY	1299	TTCCAGCCGCTCAGCATCAAGCCAGCAAGACAGACCAACCCCATCTGCT- --CAAC	1352
Db	1677	TTCCACCCCTGTCAACTTCAAGACTCAGAGGCCCAACACCGCAGCTCTTCACCCGCTCC	1736

QY 2349 AGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGCGAGGTGGAGGGGACGGG 2408
Db |||||
2814 AGCCTGGCCTCCACGGAGTTACCTGTAAAGTCTGCGTGGCGAGGTGGAAGGGGAGGC 2873
QY 2409 CAGAGCTTCAGCATCACTTCAACATCAC--AAGGACACAAGGTTTGCTGAGCTGGTG 2465
Db |||||
2874 CAGATTTTCCAGCTGTCACACAACTGTGGCCGAGACGCTGCTGGCTCCCTGGATGCTCTC 2933
QY 2466 GCTCTGAGAGTGAAGCGGGGTCCTCAGCCCTGTGTGGSCCCAGTGCCTTCAAGATCCCC 2525
Db TGCTCTGCCCGGCAATGCCATCACACCAGCTGGAGCCCTATGCCCTTCAAGATACCC 2993
QY 2526 TTCCTCATTCGGCAGAGATAAATTTCCAGCTGGACCCACCTGTAGCGGGGTGCCGAC 2585
Db CTGTCCATCGCCAAAAGATCTGCAGCAGCTGTGACGCCCCGACTCCCGGGGCAACGAC 3053
QY 2586 TGGGGGACTCTGGCCAGAACTCCACTGACAGACCATCTGACCTTCTTTGGCTTCCAAG 2645
Db |||||
3054 TGGAGGCTGTGGCGCAGAGCTGTCCATGACCGGTACCTAAACTACTTGCACCACAAA 3113
QY 2646 CCCAGCCCCACAGCCATGATCCTCAACTGTGGAGGCGGGGACCTTCCCCAAGCGCAAC 2705
Db |||||
3114 GCTAGTCCACAGGTGTCATCTTAGACCTCTGGGAGCTCGGCAACAGGATGACGGGAC 3173
QY 2706 CTGAGCAGCTGGTCAGCAGTGGCTGGACTGGGCCAGCAGACGCTGGCTCTTCACA 2765
Db |||||
3174 CTCACAGCCTGGCCAGTGGCTTGGAGGAGATGGCCAGAGTGAGATGCTGGTAGCCATG 3233
QY 2766 GTGTGGAGGCTGAGTGTCTGAG 2787
Db |||||
3234 GCCACAGATGGGATGTCTGAG 3255

RESULT 5
AY411747
LOCUS
DEFINITION Homo sapiens HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
AY411747
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
Source
gene
ORIGIN

AY411747 2532 bp DNA linear GSS 12-DEC-2003
Homo sapiens HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
AY411747
AY411747.1 GI:39767715
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2532)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 2532)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1..2532
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>2532
/locus_tag="HCM4327"

Db	1030	ACAGCCAGCGCGGCATCTACCGCGGACCGGTGATGCTCGAGACTCCACCGACAAA	1089
Qy	1428	TTCCAGCTCACCAATGGGCACCTGCTCAGGCCG	1460
Db	1090	ATCCCATGACCAACTCTCTCTGCTGGACCCCTTACCAGGCTTAAGGTCAAGGTCTTAC	1149
Qy	1461	-----CTGGGTGGC	1469
Db	1150	AGCTCCAGCACACAGGGCTCTGGCCGAGCGCTGGCAGATGGGGCTGACCTGCTGGGGTTC	1209
Qy	1470	GGCGGCGACACACTGCACCAAGCTCTCCACACTCTTGAGGCGCGAGGAGTTCGTCTCCCGC	1529
Db	1210	TTGGCGCTGGCACATACCTTAGGATTTGCGCCGGGACACCCACTTCTCGACCTGGCGC	1269
Qy	1530	CTCTCCACCCAGAACTAC-----TTCCGCTCCCTGCCCGGAGGCACACGACACATGACC	1583
Db	1270	AGCGCAGCCTCGGTTCCAGCAGCTCTTGGGCTGCCCGGAGACCCAGGAGACGCGTC	1329
Qy	1584	TATGGGACCTTCAACTTCCTCGGGGCGGCTGATATCCCTAAATACAGGTATCAGCCTC	1643
Db	1330	AGCGCACCTTTGGCTGCCCTGGGTGGAGGCTCAGCATCCCGGACACAGGGTTCAGCTTG	1389
Qy	1644	CTCATCCCCCAGATGCATACCCGAGGGAAGATCTATGATCTACCTCACGCTGCAC	1703
Db	1390	CTGGTGCCCAATGGAGCCATCCCGAGGCAAGTTCTACGAGATGATCTACTCATCAAC	1449
Qy	1704	AAGCGGAAAGCTGAGGTTGCCCTAGCTGGCTGTGACACCTGCTGAGTCCCATCGTT	1763
Db	1450	AAGCGAAGATACCTCTCCGCTTTCAGAAGGAGCCACAGACGATATTGAGCCCTCGGTG	1509
Qy	1764	AGCTGTGGACCCCTGGCGTCTGCTCACCGGCCAGTCACTCTGCTATGAGCACTGT	1823
Db	1510	ACCTGTGGACCCACAGGCTCTCTGTGCGGCCGCTATCTCTACATGCCCCACTGT	1569
Qy	1824	GGGAGGCCAGCCCTGCACAGCTGGAGCCTCGGCTCAAAAAGCAGTGTGCGAGGCGAC	1883
Db	1570	GCCGAAGTCAGTGCCCGTGACTGGATCTTTCAGTCAAGACCCAGGCCACACAGGGCCAC	1629
Qy	1884	TGGAGCAGGATGTGCTGCACCTGGGCGAGAGGCGCCCTCCACCTCTACTACTGCCAG	1943
Db	1630	TGGG---AGGAGTGGTGACCCCTGGATGAGGAGACCTGAAACACACCTGTACTGCCAG	1686
Qy	1944	CTGGAGGCAGTGCCTGACTGCTTCCACGAGCAGCTGGCGCGCTTGGCCCTGGTGGGA	2003
Db	1687	CTGGAGGCCAGGGCTGTCATCTCTGTCGACAGCTGGGACCTACGTGTTTTCAGGGC	1746
Qy	2004	GAGGCCCTCAGCTGGGTGCCCGCAAGCGCTCAAGCTGCTTCTGTTTGGCGCGGTGGCC	2063
Db	1747	GAGTCTTATTCGCTCAGCAGTCAAGGGGCTCCAGCTGGCGCTTCTTCGCCCGCCCTC	1806
Qy	2064	TGCACCTCCCTCGAGTACAAATCGGGGTCTATGCTCTGCATGACACCCACGATGACTC	2123
Db	1807	TGCACCTCCCTCGAGTACAGCCTCGGGTCTACTGCTGGAGGACACGCTGTAGCACTG	1866
Qy	2124	AAGAGGTGGTGCAGCTGGAGAACAGCTGGGGGACAGCTGATCCAGGAGCCACGGCTC	2183
Db	1867	AAGAGGTGCTGGAGCTGGAGCGGACTCTGGCGGNTACTTGGTGGAGGAGCGGAACCG	1926
Qy	2184	CTGCACCTTCAAGGACAGTTACCAAACTGCGCCTATCCATCCACGATGTGCCGAGCTCC	2243
Db	1927	CTAATGTTCAAGGACAGTTACCAAACTGCGCTCTCCCTCCATGACCTCCCGCATGCC	1986
Qy	2244	CTGTGGAAGATGAGTCTCTGTGAGCTTACAGAGATCCCTTTTATCACTCTGCAAT	2303
Db	1987	CATTGGAGGAGCAAGCTCTGGCCAAATACCAGAGATTCGCTTCTATCACTTTGGAGT	2046
Qy	2304	GGCAGCAGCGGTACTTGCACCTGCACCTTCACTCTGGAGCGTGTGACGCCCCAGCACTAGT	2363
Db	2047	GGCAGCAGAGGCCCTCCACTGCACTTTTCACTCTGGAGGAGGACAGCTTGGCTCCACA	2106
Qy	2364	GACCTTGGCTGAACTGTGGGTGTGGCAGGTGGAGGGCGACGGGACAGCTTCAGCATC	2423

Db	2107	GAGCTCACCTGCAAGATCTCTGCTGGCGGCAAGTGGAGAGGGGAGCGCCAGATATTCACGCTG	2166
Qy	2424	AACTTCAACATACACCAAG--GACACAAAGGTTTGTCTGAGCTGTGGCTCTGGAGAGTGAA	2480
Db	2167	CATACCACTCTGCAGAGACACCTGTCTGGCTCCCTGGACACTCTCTGCTGCCCCCTGGC	2226
Qy	2481	CGCGGGGTCCAGCCCTGTGTGGGCCCCAGTGCCTTCAAGATGCCCTTCCCTCATTCGGCAG	2540
Db	2227	AGCACTGTCCACCAACCAGCTGGGACCTTATGCTTCAAGATGCCACTGTCCATCCGCCAG	2286
Qy	2541	AAGATAATTCACGCTGGACCCACCTGTAGGCGGGGTCCGACTGGCGGACTCTTGGCC	2600
Db	2287	AAGATATGCAACAGCTTAGATGCCCCCACTCACGGGGCAATGACTGGCGATGTTAGCA	2346
Qy	2601	CAGAACTCCACCTGGACAGCCATCTCAGCTTCTTTTGCCTCCAAAGCCAGCCCCACAGCC	2660
Db	2347	CAGAGCTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAGAGCGCCACCGGT	2406
Qy	2661	ATGATCCTCAACTGTGTGGAGGCGGCACCTTCCCAACGGCAACCTAGCCAGCTGGCT	2720
Db	2407	GTGATCTGACCTCTGGGAAGCTCTGCAGCAGCAGCATGGGACCTCAACAGCTTGGCG	2466
Qy	2721	GCAGCAGTGGCTGGACTGGCCGACCCAGCAGCTGGCCCTCTTCACAGTGTGGAGGCTTGAG	2780
Db	2467	AGTGCCTTGGAGGAGATGGGCAAGATGAGATGCTGGTGGCTGTGGCCACCGACGGGAC	2526
Qy	2781	TGCTGA 2786	
Db	2527	TGCTGA 2532	
RESULT 6			
LOCUS	AY406492	2802 bp DNA linear	GSS 15-DEC-2003
DEFINITION	Pan troglodytes UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.		
ACCESSION	AY406492		
VERSION	AY406492.1	GI:39762466	
KEYWORDS	GSS.		
SOURCE	Pan troglodytes (chimpanzee)		
ORGANISM	Pan troglodytes		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.		
AUTHORS	Clark,A.G., Glanowski,D.S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Inferred nonneutral evolution from human-chimp-mouse orthologous gene trios		
JOURNAL	Science 302 (5652), 1960-1963 (2003)		
PUBMED	14671302		
REFERENCE	2 (bases 1 to 2802)		
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.		
FEATURES	Location/Qualifiers		
source	1..2802		
	/organism="Pan troglodytes"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9598"		
gene	<1..>2802		
	/gene="UNC5C"		
	/locus_tag="HCW2575"		
ORIGIN			
Query Match	26.5%	Score 763.2;	DB 29; Length 2802;

Best Local Similarity 52.9%; Pred. No. 8e-118; Matches 1456; Conservative 0; Mismatches 1218; Indels 78; Gaps 7;			
QY	98	GCCCGGCTGTGGCCAGCGCTCCCTGGGATAGTCTCGCCGCTGGCTCGCGGCTCGGG	157
Db	56	GCTCGTGATCTGCGCTCGCCCTGCTCAGCGCCAGCGGCACTGGCTCGCGCCCAAGA	125
QY	158	TGCCAGCAGAGTGCCACCGTGGCCCAACCCAGTGCCTGGTGCACACCCGACCTGCTTCC	217
Db	126	TGATGACTTTTTCATGAATCCCGAAGAACTTTTCCTTCGGATCCACCTGAGCCTCTGCC	185
QY	218	CCACTTCTGTGGAGCCGAGGAGTGTAATCGTCAAGAACCAAGCCAGTGTCTGTGT	277
Db	186	ACATTTCTTATTGAACCTGAGGAGCTTATATTGTGAAGATAAGCCGTGTGAACCTGTA	245
QY	278	GTGCAAGGCC-----GTGCGCGCCACGACAGATCTTCTCAAGTGCAACGGGGAGTGGT	331
Db	246	CTGTAAAGCAAGCCCTGNN	305
QY	332	GCGCCAGGTGGACACGCTGATCGAGCGCAGCACAGACGGGAGCAGTGTGAGCCGACAT	391
Db	306	NN	365
QY	392	GGAGTCCGATTAATGTCTCAAGCGCAGAGGTGAGAAAGTGTTCGGGCTGGAGGAATA	451
Db	366	NN	425
QY	452	CTGTGCGCAGTGTGCGATGGAGCTCTCGGGACCAACAGAGTCAGAAAGCCTACAT	511
Db	426	NN	485
QY	512	CCGATAGCCAGATTGCGCAAGAACTTCGAGCAGAGCGCGCTGGCCAGGAGTGTCCCT	571
Db	486	NN	545
QY	572	GGAGCAGGCGATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	631
Db	546	GGAACAGGAAGTCTTACTCCAGTGTGCGCCACCTGAAGGATCCAGTGGCTGAGNNNN	605
QY	632	GTGGCTCCGGAACGAGGACCTGTGGACCGCTCCCTGGACCCCAATGATATACATCAGCG	691
Db	606	NNNNNGAATAATGAAGACATAATGATCCGTTGAAGATCGGAATTTTATATATAT	665
QY	692	GGAGCAGCGCTGTGTGTGCGACAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	751
Db	666	TGATCAACACTCATATAAGCAGCGCCGACTCTGTGATCTGCAATTTATACCTTGT	725
QY	752	GGCCAGAGACATCGTGGCAGCTGCGCGAGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTG	811
Db	726	TGCCAAAACATTTGTGCCAAGAGAAAGTACAACTGCCACTGTATATGCTATGTC	785
QY	812	CGGTGGGTGTGACGTGGACCGAGTGGTCCGTCTGCGAGCCAGCTGTGGGCGCGGCTG	871
Db	786	CGGTGGCTGTCCCTGACCGAGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	845
QY	872	GCAGAACCGGAGCTGACCAACCGCGCGCTCTCAACGGCGCGCTTCTGTGTGA	931
Db	846	TCAGAAACGTACAAGACTTGTACCAACCGGACCACTCAATGGGGGTGCTTCTGTGA	905
QY	932	GGGGCAGATGTCCATGACCGCACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	991
Db	906	AGGGCAGAGTGTGCAAA---AATAGCTGTACTACGTATTGCCAGTGGATGGCAGGTG	962
QY	992	GAGCCGTGGAGCAAGTGTGCGCTGTGGCTGGAGTGCACCCACTGGCGGAGCGGTGA	1051
Db	963	GAGCCATGGAGCAAGTGTGTACTTGTGAACTGAGTGCACCCACTGGCGCAGGAGGA	1022
QY	1052	GTGCTGTGACCCAGCACCCGCAACCGAGGGAGGAGTGCCAGGCACTGACTGACAC	1111
Db	1023	GTGACGGCGCAGCCCCCAAGNATGAGGACAGGACTGCGATGGCTCGTCTTGCAATC	1082
QY	1112	CCGAACTGTACAGTGAACCTGTGTGTACACAGTGTCTGTGCTGTGAGGAGCTGGCCCT	1171
Db			
1083	CAAGNACTGCACTGATGGGCTTTTGATGCAGACTGCTCCTGATTCAGATGATGTTGCTCT	1142	
QY	1172	CTATGTGGG---CCTCATGCGCGTGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1228
Db	1143	CTATGTTGGGATTTGTATAGCAGTATGTTTGGCTGGGATCTCTGTAGTGTGGGCTT	1202
QY	1229	CCTGTTTATTTCCGGAAGAAGGAGGCTGAGCTCAGATGTGGCTGACTGCTGCTGCTGCT	1288
Db	1203	GTTTGTGTATCGGAAGAATCATCTGACTTTTGTAGTCAGATATTATTGACTCTTGGCACT	1262
QY	1289	CACCTCAGGCTTTCAGCGCGTCAAGCCAGCCAGCAAGAAAGACAGACACCCCATCTGCT	1348
Db	1263	CAATGGGGGCTTTTCAAGCTGTGAACATCAAG-----GCAGCAAGACAAGATCTGCT	1313
QY	1349	CACCATCAAGCGGACCTCAGCACCAACCACTAGCAGGGCAGTCTCTGCCCGGCA	1408
Db	1314	GGCTGTACCCCGCAGACCTCAGCTCAGCTGAGCCATGTACAGAGCACTGTCTATGCCCT	1373
QY	1409	GGATGGGCCAGCCCAAGTTCAGCTCAACAAATGGGACCTGTCTAGCCCCCTGGGTGG	1468
Db	1374	GCATGACGTCTCAGACAAATCCCAATGACCAACTCTCCAATCTGGATCCACTGCCAA	1433
QY	1469	CGGCGCCACACACTGCACACACAGCT-----CTCCCACTCTGAGGCGCAGGA	1516
Db	1434	CCTGAAATCAAGTGTACAAACCTCAGGTGTGTCACCCCCAAGATGACCTCTCTGA	1493
QY	1517	GTTGCTCTCCGCTCTCCACCCAGAACTACTTCCGCTCCTGCTGCCCGCAGCACACGAA	1576
Db	1494	GTTTACGTCAAGTGTCCCTCAGATGACCCAGTCAITGTTGGAGATGAAGCCCTCAG	1553
QY	1577	CATGA-----CCTATGGGACCTT	1594
Db	1554	CCTGAAGAACAGAGCTAGCAAGGAGCAGTGCATCTCTGTACCGCAITTTGGCAGCTT	1613
QY	1595	CAACTTCTCGGGCGCGCTGATGATCCCTAATACAGGTATCAGCTCTCTCATCCCCC	1654
Db	1614	CNNNTGCTGGNNNNACCTTATTTGTTCCCAATCAGGAGTCAGCTGTGATTTCCCGC	1673
QY	1655	AGATGCCATACCCCGAGGGAAGATCTATGAGATCTACTCACGCTGCACAAAGCGGAGA	1714
Db	1674	TGGGGCAITTTCCCAAGGAGAGTCTACGAAATGTATGTACTGTACAGAGAAAGAAAC	1733
QY	1715	CGTGAAGTTCCTCCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1774
Db	1734	TATGAGGCCACCTCGATGACTCTCAGACATTTTGACCCCTGTGTGAGTGTGGGC	1793
QY	1775	CCCTGCGCTCTGCTCACCGCGCAGTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1834
Db	1794	CCAGAGCTCTGCTCACCGCGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1853
QY	1835	CCCTGACAGCTGAGCTCGGCTCAAAAGCAGTCTGTCAGGGCAGCTGTCAGGAGCAGGA	1894
Db	1854	TACCGAGGACTGGAAATATCTGCTCAAGAACCCAGGACAGCAGGAGCAGTGGG---AGGA	1910
QY	1895	TGCTGCTGCACTGGCGAGGAGCGCCCTCCACCTCTACTCTACTGCTGCTGCTGCTGCTG	1954
Db	1911	TGTGTGTGTGTGGGAGGAGAACTTCAACCCCTCTCTACTATTCAGTGTGATCAG	1970
QY	1955	TGCTGCTACGTCTTTCACCGAGCAGTGGCGCTTTCCTGCTGCTGCTGCTGCTGCTGCTG	2014
Db	1971	GGCTTGCACATCTCTCAGAGAACTCAGCAGCTTTCAGCCCTGGTAGGACATTTCCACCAC	2030
QY	2015	CGTGTGCTCCGCAAGCGCTCAAGCTGCTTCTGTTTGGCGCGGTGGCTGCTGCTGCTGCTG	2074
Db	2031	CAAGCGGTGCGAAGCGCTCAAGCTGGCCATCTTTGGGCGCTGCTGCTGCTGCTGCTGCTG	2090
QY	2075	CGAGTACAACTCCGGGTCTACTGCTGATGACACCAACGATGACACTCAAGAGGTGT	2134
Db	2091	GGAGTACAGCATCCGAGTCTACTGTCTGATGACACCCAGGATGCCCTGAAGAAATTTT	2150
QY	2135	GCAGCTGAGAGAGCTGGGGGACAGTATCCAGAGCCACCGGTCTCTGACTTCAA	2194
Db	2151	ACATCTTGAGAGACAGATGGGAGGACGCTCCTAGAAGAACCTAAGGCTCTTCTTTTAA	2210


```
QY 2195 GGACAGTATACCAACACTGGCGCTATCCATCCAGATGTGCCAGCTCCCTGTGGAGAG 2254
DB |||||
DB 2211 AGCAGACACCAACACTGGCGCTGTCAATTCAGATATGCCCATTCCTCTGGAAGAG 2270
QY |||||
DB 2255 TAAGCTCCCTGTGAGTACAGAGATCCCTTTTATCATCTGGAATGGCAGCGAGCG 2314
DB |||||
DB 2271 CAAATGCTGGCTAAATATCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2330
QY |||||
DB 2315 GTACTTGCACGTGACCTTCACTCTGGAAGATTAGCTGAACACAGTGGAGCTGTTG 2374
DB |||||
DB 2331 NAACCTGCATGACCTTCACTCTGGAAGATTAGCTGAACACAGTGGAGCTGTTG 2390
QY |||||
DB 2375 CAAAGCTGTGGGTGGAGGTGAGGGCGAGCGGCGAGAGCTTCAGCATCACTTCAACAT 2434
DB |||||
DB 2391 CAAACTCTGTGTGGCGCAGGTGGAAGGAGGAGGAGAGCTTCAGCTCAACTGCACGT 2450
QY |||||
DB 2435 CACCAAGGACACAGGTTTCTGAGCTGTGGCTCTGAGAGTGAAGCGGGGTCCCGAGC 2494
DB |||||
DB 2451 GTGAGAGGAACCTACTGGCATCGATTTCGCTGCTGGATCTCTGCGAACCATCACAC 2510
QY |||||
DB 2495 CCTGGTGGGCGCCAGTCCCTTCAAGATCCCTTCTCATTTGGGAGAGATAATTTCCAG 2554
DB |||||
DB 2511 GGTACGGGGCCAGTCTTTTACATCCCTCTCCCTATCCGGCAGAGAGCTCTGTAGCAG 2570
QY |||||
DB 2555 CCTGGAGCCACCTGTAGGGGGGTGGCGAGTGGCGACTCTGGCCAGAGAACTCCACCT 2614
DB |||||
DB 2571 CCTGGATGCCCCCAGACGAGAGGCCATGACTGAGAGATGTGCCCCATAAGCANNNNNN 2630
QY |||||
DB 2615 GGACAGCATCTCAGCTTTTGTCTCAGCCCGAGCCAGCCATGATCTCTCAACCT 2674
DB |||||
DB 2631 NNNNNGTACTTGAAATTTTGGCCAAATCCAGCCCACTGGCGTAACTCTGGATCT 2690
QY |||||
DB 2675 GTGGAGCGCGGCACCTTCCCAACGGCAACCTCAGCCAGCTGGCTGCGAGAGTGGCTGG 2734
DB |||||
DB 2691 TTGGGAAGCAGAACTTCCAGATGGAACCTCAGCATGTGGCAGCTGCTTGGGAAGA 2750
QY |||||
DB 2735 ACTGGGCGAGCAGCTGCGCTTTTCAAGTGTGGAGCTGAGTGCCTCA 2786
DB |||||
DB 2751 AATGGGAAGACATGAAAGCGTGTCTTTAGCAGCAGAGGGCAGTATTAA 2802
```

```
RESULT 7
BI758231
LOCUS 603029876P1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200171 5',
DEFINITION mRNA sequence.
ACCESSION BI758231
VERSION BI758231.1 GI:15749809
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1034)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgepbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11501 row: g column: 20
High quality sequence stop: 793.
Location/Qualifiers
1..1034
/organism="Homo sapiens"
/mol_type="mRNA"
```

/db_xref="taxon:9606"									
/clone="IMAGE:5200171"									
/lab_host="DH10B"									
/clone_lib="NIH_MGC_114"									
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC Library."									
ORIGIN									
Query Match 26.2%; Score 756.2; DB 12; Length 1034;									
Best Local Similarity 89.7%; Pred. No. 7.5e-117;									
Matches 905; Conservative 0; Mismatches 88; Indels 16; Gaps 8;									
QY	191	GCCTGGTGCACCCGACCTGCTTCCCACTTCTCTGGTGGAGCCGAGATGTGTACAT	250						
DB	1	GCCTGGTGCACCCGACCTGCTTCCCACTTCTCTGGTGGAGCCGAGATGTGTACAT	60						
QY	251	CGTCAAGAAACAAGCCAGTGTCTGTGTCGAAGCCGCTGCCCGCAGATCTTCTT	310						
DB	61	CGTCAAGAAACAAGCCAGTGTCTGTGTCGAAGCCGCTGCCCGCAGATCTTCTT	120						
QY	311	CAAGTGCACCGGAGTGGTGGCCAGGTGGACCATGATCGAGCGCAGCAGACGG	370						
DB	121	CAAGTGCACCGGAGTGGTGGCCAGGTGGACCATGATCGAGCGCAGCAGACGG	180						
QY	371	GAGCAGTGTGTGAGCCGACCATGAGGTCCGATTAATGTCTCAAGGAGCAGGTCTGAGAA	430						
DB	181	GAGCAGTGTGTGAGCCGACCATGAGGTCCGATTAATGTCTCAAGGAGCAGGTCTGAGAA	240						
QY	431	GGTGTTCGGCTGAGGAAATATCTGTGTCAGTGGTGGCATGGAGTCTCTCGGCAACCAC	490						
DB	241	GGTGTTCGGCTGAGGAAATATCTGTGTCAGTGGTGGCATGGAGTCTCTCGGCAACCAC	300						
QY	491	CAAGAGTCAGAGCCCTACATCCGATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCC	550						
DB	301	CAAGAGTCAGAGCCCTACATCCGATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCC	360						
QY	551	GCTGCCCAAGAGGTGCTCTTGGAGCAGGGCATCGTGTGCTGCTGCCCTGCCCTCCACGGAGG	610						
DB	361	GCTGCCCAAGAGGTGCTCTTGGAGCAGGGCATCGTGTGCTGCTGCCCTGCCCTCCACGGAGG	420						
QY	611	CATCCCTCCAGCCGAGGTGGAGTGGTCCGGAACGAGGACCTGTGGACCCGTCCCTCGA	670						
DB	421	CATCCCTCCAGCCGAGGTGGAGTGGTCCGGAACGAGGACCTGTGGACCCGTCCCTCGA	480						
QY	671	CCCCAATGTATACATCAGCGGGAGCAGCCTGCTGTGTGGCAGAGGCCCTTGTCTGA	730						
DB	481	CCCCAATGTATACATCAGCGGGAGCAGCCTGCTGTGTGGCAGAGGCCCTTGTCTGA	540						
QY	731	CACGCCCAACTACACCTGCTGGCAGAAACATCGTGGCAGCTGCCCGCAGGCTCCGC	790						
DB	541	CACGCCCAACTACACCTGCTGGCAGAAACATCGTGGCAGCTGCCCGCAGGCTCCGC	600						
QY	791	TGCTGTATCTCTACGTGAACCGTGGTGTGCA-CGTGGACCGAGTGGTCCCTCTGCA	849						
DB	601	TGCTGTATCTCTACGTGAACCGTGGTGTGCA-CGTGGACCGAGTGGTCCCTCTGCA	659						
QY	850	GCGCCAGCTGGCGGGCTGGCAGAAACCGAGCCGAGCTGCACCAACCCGCGGCTC	909						
DB	660	GCGCCAGCTGGCGGGCTGGCAGAAACCGAGCCGAGCTGCACCAACCCGCGGCTC	719						
QY	910	TCAACGGGGGGCTTTTGTGAGGGGAGAAATGTCCATGACCGCACCGTCTCTCTGC	969						
DB	720	TCAACGGGGGGCTTTTGTGAGGGGAGAAATGTCCATGACCGCACCGTCTCTCTGC	774						
QY	970	TGTGTCTGTGAGCGGAGCTGGAGCCCGTGGAGCAAGTGTGCGGCTGT--GGCTGA	1027						

```
Db 775 CTGTGCCAGTGGACGGCAGCTGTAGACAGTGGAGCCAGTGGTGGCCCTGTTGGGCTTGGC 834
Qy 1028 CTGACCCCACT--GGCGAGCGCTGAGTCTCTGACCCAGCAGCCCGCAAGCGGGGAG 1085
Db 835 TTGACCCCACTTGGCGGAGCGGAGTCTCTGAACCCAGCAGCCCGGCAAGCGGGGG 894
Qy 1086 GAGTG---CCAGGCACTGACCTGGACACCGCCCAACTGTACAGTGAACCTCTGTGTACAC 1142
Db 895 AGGTGTGCCAGTCACTGGACCTGGACCCCGGA-TGGTCAAGTGAAGTCTGTGT-CCC 952
Qy 1143 AGTCTTCTGGCCCTGAGGAGTGGCCCTCTATGTGGGCCCTCATCGCG 1191
Db 953 ACGGTTCTGGCCCTGAGGACTTGGCCCTCTATGTGGGCCCTCATCCCG 1001

RESULT 8
AY411749
LOCUS 2532 bp DNA linear GSS 12-DEC-2003
DEFINITION Mus musculus HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY411749
VERSION AY411749.1 GI:39767717
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2532)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2532)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..2532
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2532
/locus_tag="HCM4327"

gene
ORIGIN
Query Match 25.2%; Score 726.4; DB 29; Length 2532;
Best Local Similarity 59.1%; Pred. No. 1.2e-111;
Matches 1492; Conservative 0; Mismatches 896; Indels 138; Gaps 9;

Qy 393 GAGTCCGATTAATGTTCTCAGGCGAGCGGTGCGAAGGTGTTGGGCTGGAGGAATAC 452
Db 13 GAGGTGCAGATGAGGTGTCAAGGCAAGTGGAGGAAGTCTTGGGCTCGAGGACTAC 72
Qy 453 TGGTGCAGTGGCTGCATGGAGTCTCTGGGACCCAGCAAGTCAAGGCTCATATC 512
Db 73 TGGTGCAGTGGCTGGAGTCTTTCGGAACTTACCAAGTCCGCGAGCTATATC 132
Qy 513 CGCATAGCCAGATTGGCGAAGAACTTCGAGCAGGAGCGCTGGCCAGGAGTGTCCCTG 572
Db 133 CGCATTGCCTACTTGGCGAAGAACTTTCACGAGGAGCTCTTGCCAGGAGGTACCTTG 192
Qy 573 GAGCAGGCGCATCGTGTGCCCTGCCGCTCCAGCGGAGGATCCCTCCAGCCGAGGTGGAG 632
Db 1270 AGTGCCAGCCTTGGTTCACAGCACCTCTCTGGGCTTACCTCGGAGCCCGCAGCAGTGTGTC 1329
```

```
Db 193 GATCATGAGGTCTCTTGTGAGTGGCGGCCCAACCGGAGGAGTGCCTGTGGCTGAGGTGAA 252
Qy 633 TGGTTCGGGAACGAGGACCTTGGTGGACCCGTCCTTGGACCCCAATGTATACATCACGCGG 692
Db 253 TGGCTCAAGAAATGAAGATGTCAATTGACCCGCTCAGGACACTAATCTCTGCTCAACATT 312
Qy 693 GAGCAGACCTGTTGGTGGCAGACGCGCGCTTGTGACACGCGCAACTACACTGCGTG 752
Db 313 GACCACAACCTTCATCATCCGACAGGCGCTCTCAGACACGCGCAACTACACTGTGTG 372
Qy 753 GCCAAGAAATCTGTGGCAGCTCGCGCAGGCGCTCCGCTGCTGTCTCATCTGTCTACGTGAA 812
Db 373 GCCAAGAAATCTGTGGCAGGCGCGGAGACACCGGCCACAGTCATCGTCTATGTGAAT 432
Qy 813 GGTGGTGGTTCGACGTGGACCGAGTGGTCTGTGACGCGCCAGCTGTGGCGCGCGCTGG 872
Db 433 GGAGGCTGGTCCAGCTGGCAGAGTGGTCAACCTGTTCCAAATCGTGTGGCCGAGGCTGG 492
Qy 873 CAGAAACGAGCGGAGCTGACCAACCCGCGGCTCTCAACGCGGCGCTTCTGTGTGAG 932
Db 493 CAGAAGCGTACTCGGACCTGCACCAATCCAGCCCCACTCAATGGAGCGCGCTTCTGTGAG 552
Qy 933 GGCAGAAATGTCCATGACCGCACCGTCTCTCTCTGTCTGTCTGTGTCGTGGACGCGAGTGG 992
Db 553 GGACAGGCTTCCAGAA--GACAGTGTGACACCGTGTGCCAGTGGATGGAGCGTGG 609
Qy 993 AGCCGCTGGAGCAAGTGGTGGCGCTGTGGGCTGGACTGACCCCACTGGCGGAGCCGTGAG 1052
Db 610 ACCGAGTGGAGCAAGTGGTGTGCTGCTGACGACAGAGTGTGCGCACTGGCGCAGCGCGAG 669
Qy 1053 TGTCTGTACCCACACCCCGCAACGAGGAGGAGTGCACGAGGCACTGACCTGGACACC 1112
Db 670 TGCATGGCACCGCCACCCCAAGCGGAGCGCTGACTGACAGCGGAGCGCTACTTGAATCC 729
Qy 1113 CGCAACTGTACCACTGACCTCTGTGTACACAGTCTTCTGCGCTGAGGAC----- 1163
Db 730 AAGAACTGCACTGATGGGCTGTGCTGTGCTGTGAATCAGAGAACTCTAAGACACCTTAAAGC 789
Qy 1164 -----GTGGCCCTCTATGTGGGCTCATGTGGGCTTCATGCGCGTGGCGCTC 1199
Db 790 CACCCCTGGAGACATCGGGAGATGTGGCACTGTACGAGGCTTGTGGTGGCGCTTTT 849
Qy 1200 TGCTCTGCTCTGCTGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1256
Db 850 GTGGTGTGAGCGTCTCTATGCGCGTGGAGTGTATATACCGGAGAACTCCCGGAG 909
Qy 1257 CTGGACTCAGATGTGGCTGACTCGT---CCATTCTCACTCAGGCTTCCAGCCGCTCAGC 1313
Db 910 TTGACACCGACATCACCGACTCTCTGCGGCGCTCCTGCTGGTGGCTTCCACCTGTCAAC 969
Qy 1314 ATCAAGCCAGCAAGACAGACAAACCCCATCTGCT-----CACCATCCAGCGGACCTC 1367
Db 970 TTCAAGACTGCAAGGCCCAACCAACCGCAGCTCTGCAACCGCTCCGCTCCAGACCTA 1029
Qy 1368 AGCACACACCACTTACAGGCGAGTCTCTCTCCCGGAGAGTGGCGGAGTGGCGGAGTGG 1427
Db 1030 ACGCGAGTGTGGCACTTACCGCGGCGCTGTGTATGCTGCGGAGTCTCCGCGGAG 1089
Qy 1428 TTCCAGCTCACCATGGGCACTGCT----- 1453
Db 1090 ATCCCATGACTAATTGCGCCCTGTGGATCCCTGCGGAGCTCAAGATCAAGTCTAT 1149
Qy 1454 -----CAGCCCCCTGGGTGGC 1469
Db 1150 AACTCCAGCACCATCGTTTCTGGGCTGTGGCTGGTGTGATGGAGCCGACCTGTGGGTGTC 1209
Qy 1470 GCGCGGCAACACTGCACCAAGCTCTCCACTCTGTAGGCGCGAGAGTGTCTCTCCGCG 1529
Db 1210 CTCGCGCGGCGGACGCTACCCAGGCAATTTCTCCGCGGACACCCATTTCTCTGCACTGCGC 1269
Qy 1530 CTCTCCACCCAGAACT-----ACTTCGCTCTCTGCGGAGGACACAGCAAGTATGACC 1583
Db 1270 AGTGCCAGCCTTGGTTCACAGCACCTCTCTGGGCTTACCTCGGAGCCCGCAGCAGTGTGTC 1329
```

QY	1584	TATGGACCTTCAACTTCTCTGGGGCCCGCTGATGATCCCTAAATACAGGTATCAGCCTC	1643
Db	1330	AGCGGCACTTTGGTTGGCTTGGGAGAAAGGCTGAGCCTCCCCGGCACAGGGGTGAGCCTG	1389
QY	1644	CTCATCCCCCAGATGCCATACCCCGAGGAAGATCTATGAGATCTACCTCAGCGTGCAC	1703
Db	1390	TTGGTACCAAAATGAGGCCATTCCTCCAGGGCAAGTTCTATGACCTGTATCTACATATCAAC	1449
QY	1704	AAGCGGAAGACGTGAGGTTGGCCCTAGCTGGTGTGACAGCCCTGTGAGTCCCATCGTT	1763
Db	1450	AAGCGGAAGACCTCCACATTTTCAGAGGTTCCACAGACAGTAITGAGCCCTCGGTG	1509
QY	1764	AGCTGTGAGACCCCTCGGCTCTGCTACCCGGCCAGTCACTCTGGCTATGAGACACTGT	1823
Db	1510	ACCTGTGGGCCACAGAGCCTACTCTGTGCGCCCTGTGCTCCTACCGTGCCTCACTGT	1569
QY	1824	GGGAGCCAGCCCTGACAGCTGGAGCTGGCTCAAAAGCAGTCTGCGAGGGGAGC	1883
Db	1570	GCTGAAGTCATCGCTGGAGACTGGATCTTTTCAGCTCAAGACCCAGGCCCATCAGGGCCAC	1629
QY	1884	TGGAGCAGAGATGTGTGACCTGGCGAGGAGCGCCCTCCACACTTCTACTCTGCTCAG	1943
Db	1630	TGGG---AGAGGTGGTGACTTTGGATGAGAGACCCTCAACACACCCTGCTACTGCTCAG	1686
QY	1944	CTGAGGCCAGTGTCTGCTAGCTTTCCAGAGAGCTGGGCGCTTTGGCCCTGGTGGGA	2003
Db	1687	CTGAGGCTAAGTCTCTGCCACATCTCTGTGACAGCTGGGTACCTACGTATTATCATGGC	1746
QY	2004	GAGGCTCAGCGTGGTGGCGGCAAGCGCTCAAGCTGCTTCTGTTGGCGGCTGGCC	2063
Db	1747	GAGTCTTACTCTGCTGTGAGTCAAGCGGCTCAGCTGGCCATCTTCGCCCCAGCCCTC	1806
QY	2064	TGCACCTCCCTCGAGTACAAATCCGGGTCTACTGCTGATGACACCCAGATGCACTC	2123
Db	1807	TGCACCTCCCTGGAGTATAGCTCAGGCTCTACTGCTGGAGGACACACCTGTAGCACTG	1866
QY	2124	AAGAGTGTGTGAGCTGGAGAACAGCTGGGGGGAGAGCTGATCCAGGAGCCAGGGTC	2183
Db	1867	AAGAGGTCTCTGGAGCTGGAGAGGACTCTGGTGGCTACTTTGGTGGAGGCCCAAGCCT	1926
QY	2184	CTGCACCTCAGGACAGTATCCACAACTGGCGCTATCTCCATCCAGATGTGCCAGTCC	2243
Db	1927	TTGCTCTTTAAGGACAGTTTACCAACTAGCTGCTCTCTCCATGATATCCCCCATGCC	1986
QY	2244	CTGTGGAAGATGATGCTCTTGTGACGCTACAGGAGATCCCTTTTATCATCTGGAAT	2303
Db	1987	CAGTGGAGGAGCAAACTACTGGCCAAAGTACAGGAGATTCCTCTTACCAAGTCTGGAAT	2046
QY	2304	GGCAGGAGGGTACTTGTGACGCTTACCTTACCTGGAGCGTGTGAGCCCGACACTAGT	2363
Db	2047	GGCAGCCAGAGCCCTGCACTGCACTTTACCTCCCTGGAGAGGATAGCCTGGCCTCCACG	2106
QY	2364	GACCTGGCTCAGAGCTGTGGGTGTGGCAGTGTGAGGGCGAGCGGACAGCTTCAGCATC	2423
Db	2107	GAGTTCACTTAAAGTCTGCTGTGGGAGAGTGTGAGGGAGGAGCCAGATTTTCCAGCTG	2166
QY	2424	AACCTTCAACATCACC---AAGGACACAGGTTTGTGAGCTGTGCTGTCTGTGAGAGTAA	2480
Db	2167	CACACAAAGTTGGCCGAGACGCTGTGGCTCCCTGGATGTCTCTGTCTGTGCCCGGCG	2226
QY	2481	CGGGGGTCCAGCGCTGGTGGGCCCCAGTGGCTTCAGATCCCTTCTCATTCGGCAG	2540
Db	2227	AATGCCATCACACCCAGCTGGGACCTTATGCTTTCAAGATACCCCTGTCTCATTCGCCAA	2286
QY	2541	AAGATAATTTCCAGCTGGACCCACCTGTAGGGCGGGTGGCGGACTGGCGACTCTGGCC	2600
Db	2287	AAGATCTGAGAGGCTGAGAGCCCCCAACTCCCGGGGCAAGCACTGGAGGCTGTGGCG	2346
QY	2601	CAGAACTCCACTGGAGAGGCACTGACGTTCTTTTGGCTTCCAAAGCCAGGCCCAAGCC	2660
Db	2347	CAGAAGCTGTCCATGAGCCGGTACCTAAACTACTTTCGCCACCAAAAGCTAGTCCCAAGGT	2406
QY	2661	ATGATCTCAACTGTGGAGCGCGCACTTCCCAAGCGCAACCTCAGCCAGCTGGCT	2720
Db	2407	GTCATTTTAGACCTCTGGGAAGCTCGGCAACAGGATGACGGGACCTCAACAGCTGGCC	2466
QY	2721	GCAGCAGTGGCTGAGCTGGGCCAGCCAGCGCTGGCTCTTTCACAGTGTGGAGGCTGAG	2780
Db	2467	AGTGCTTGGAGGAGATGGCAAGAGTGTGATGCTGTAGCCATGGCCACAGATGGCGAT	2526
QY	2781	TGCTGA 2786	
Db	2527	TGCTGA 2532	
RESULT 9			
BX364574			
LOCUS			
DEFINITION			
BX364574 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens			
cDNA clone CS0DC023YB10 5-PRIME, mRNA sequence.			
ACCESSION			
BX364574			
VERSION			
BX364574.1 GI:30368812			
KEYWORDS			
EST.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1 (bases 1 to 909)			
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.			
Full-length cDNA libraries and normalization			
Unpublished (2001)			
COMMENT			
Contact: Genoscope			
Genoscope - Centre National de Sequencage			
BP 191 91006 EVRY cedex - France			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
Library was constructed by Life Technologies, a division of			
Invitrogen. This sequence belongs to sequence cluster 3529.f For			
more information about this cluster, see			
http://www.genoscope.cns.fr/			
cgi-bin/cluster.cgi?seq=CS1AC006ZF06QP1&cluster=3529.f. Contact :			
Feng Liang Email : fliang@lifetech.com URL :			
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600			
Faraday Avenue Genoscope sequence ID : CS1AC006ZF06QP1.			
FEATURES			
Location/Qualifiers			
1..909			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="CS0DC023YB10"			
/tissue type="NEUROBLASTOMA COT 25-NORMALIZED"			
/clone lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"			
/note="1st strand cDNA was primed with a NotI-oligo (dr)			
primer. Five prime end enriched, double-strand cDNA was			
digested with Not I and cloned into the Not I and EcoR V			
sites of the pCMVSPORT 6 vector. Library was normalized."			
ORIGIN			
Query Match			
Best Local Similarity 25.2%; Score 726.2; DB 13; Length 909;			
Matches 739; Conservative 98.4%; Pred. No. 7.9e-112;			
Mismatches 4; Indels 1; Gaps 1;			
QY	1	AGCTGGGGCTCCGGGCTGAGCGCTAAAGCGCCCTCCCGCCCGGGGGCCCCGGCCCG	60
Db	160	AGCTGGGGCTCCGGGCTGAGCGCTAAAGCGCCCTCCCGCCCGGGGGCCCCGGCCCG	219
QY	61	GCCCGCGGCTCCCGCCCGCGCCGCTGGCGCTCCGCGCGGCTGTGGCGAGCGCTCC	120
Db	220	GCCCGCGGCTCCCGCCCGCGCCGCTGGCGCTCCGCGCGGCTGTGGCGAGCGCTCC	279
QY	121	TGGGCATAGTCTCGCGCTTGGCTCCGCGGCTCGGFTGCCAGAGTGCACCGTGG-180	
Db	280	TGGGCATAGTCTCGCGCTTGGCTCCGCGGCTCGGFTGCCAGAGTGCACCGTGG-339	
QY	181	CCAAACCCAGTGTGGTGCACACCGGACCTGCTTCCCCTCTCTGTGTGGAGCCCGAGG	240


```

QY 1278 TCGTCCATCTTCACCTCAGGCTTCCAGCCGTGAGCATCAAGCCAGCAAGCAGACAAC 1337
Db 1341 TCGTCCATCTTCACCTCAGGCTTCCAGCCGTGAGCATCAAGCCAGCAAGCAGACAAC 1400
QY 1338 CCCCATCTGTCTACCATCCAGCCGGAGCTCAG---CACCACCCACCACTTACCCAGGGCAGT 1394
Db 1401 CCCCATCTGTCTACCATCCAGCCGGAGCTCAGCAACCAACCACTTACCCAGGGCAGT 1460
QY 1395 CTCGTCTCCCGGAGGATGGCCCGAGCCCAAGTTCAGCTTCCCAATGGCCACTCTCTC 1454
Db 1461 CTCGTCTCCCGGAGGATGGCCCGAGCCCAAGTTCAGCTTCCCAATGGCCACTCTCTC 1520
QY 1455 AGCCCCCTGGGT 1466
Db 1521 AGCCCCCTGGGT 1532

RESULT 12
LOCUS B1818609 818 bp mRNA linear EST 04-OCT-2001
DEFINITION 601033362F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174559 5',
mRNA sequence.
ACCESSION B1818609.1 GI:15929902
VERSION B1818609
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rcapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11434 row: 1 column: 16
High quality sequence stop: 744.
FEATURES
Source Location/Qualifiers
1..818
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5174559"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/notes="Organ: pooled brain, lung, testis; Vector:
pCMV-SpOR6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcorV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 24.6%; Score 710; DB 12; Length 818;
Best Local Similarity 97.1%; Pred. No. 4e-109;
Matches 744; Conservative 0; Mismatches 20; Indels 2; Gaps 2;

QY 84 GCCATGCCCTCCGGCCCGGCTGTGGCCAGCGCTCTCGGCATGCTTCGGCGCTTGG 143
Db 1 GCCATGCCCTCCGGCCCGGCTGTGGCCAGCGCTCTCGGCATGCTTCCTCGCGCTTGG 60
QY 144 CTCGCGGGCTCGGGTGGCCAGCAGAGTGCCACCGCTGGCCCAACCCAGTGCTGTGGTCCAAAC 203

```

```

Db 61 CTCGCGGGCTCGGGTGGCCAGCAGAGTGCCACCGTGGCCAAACCCAGTGCTGTGGTCCAAAC 120
QY 204 CCGGACCTGTTTCCCACTTCTCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAAACAAG 263
Db 121 CCGGACCTGTTTCCCACTTCTCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAAACAAG 180
QY 264 CCAGTGTCTGTGTGTCAGAGGGCTGCCCGCCACGAGATCTTCTTCAAGTGAACAGGG 323
Db 181 CCAGTGTCTGTGTGTCAGAGGGCTGCCCGCCACGAGATCTTCTTCAAGTGAACAGGG 240
QY 324 GAGTGGGTGGCCAGAGTGGACCACTGATCGAGCGCAGCACAGACGGAGCAGTGTGTGAG 383
Db 241 GAGTGGGTGGCCAGAGTGGACCACTGATCGAGCGCAGCACAGACGGAGCAGTGTGTGAG 300
QY 384 CCGACCATGAGGTCCGCATTAAATGTCTCAAGGCAGCAGGTTCGAGAAAGTGTTCGGGCTG 443
Db 301 CCCACCATGAGGTCCGCATTAAATGTCTCAAGGCAGCAGGTTCGAGAAAGTGTTCGGGCTG 360
QY 444 GAGGATACTGGTGGCCAGTGGGATGGAGCTCTCGGGCACCACCAAGAGTCAAGAG 503
Db 361 GAGGATACTGGTGGCCAGTGGGATGGAGCTCTCGGGCACCACCAAGAGTCAAGAG 420
QY 504 GCCTTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAAGGAG 563
Db 421 GCCTTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAAGGAG 480
QY 564 GTGTCCCTGGAGCAGGGCATCGTGTCTGCCCTGCCCTCCACCGAGGGGCATCCCTCCAGCC 623
Db 481 GTGTCCCTGGAGCAGGGCATCGTGTCTGCCCTGCCCTCCACCGAGGGGCATCCCTCCAGCC 540
QY 624 GAGTGGAGTGGCTCCGGAAACGAGGACCTGTGTGGACCCCTCCCTGGACCCCAATGTATAC 683
Db 541 GAGTGGAGTGGCTCCGGAAACGAGGACCTGTGTGGACCCCTCCCTGGACCCCAATGTATAC 600
QY 684 ATACCGCGGAGCAGACAGCCCTGTGTGCGACAGCCCGCTGTGTGACACGGCCAACTAC 743
Db 601 ATACCGCGGAGCAGACAGCCCTGTGTGCGACAGCCCGCTGTGTGACACGGCCAACTAC 660
QY 744 ACTGTGCTGCCCAAGAACATCGTGGCAAGTTC-GCCGACAGCGCTCCGCTGCTGTCAFCGT 802
Db 661 ACCTGTGCTGCCCAAGAACATCGTGGCAGCTGACCGCGAGGGGCTCCGCTGCTGTCAFCGT 720
QY 803 CTACGTGAACGGTGGTGGTGGAGC-TGGACCGAGTGGTCCGCTG 847
Db 721 CTACGTGGACGGCAGCTGGAGCCCGCTGGAGCCAGTGGTCCGGCTG 766

```

```

RESULT 13
LOCUS BX348193/c 843 bp mRNA linear EST 05-MAY-2003
DEFINITION BX348193 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
cDNA clone CS0DB008YE02 5-PRIME, mRNA sequence.
ACCESSION BX348193
VERSION BX348193.1 GI:30367258
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 843)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3239.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAF004ZD01_AF00293_1&cluster=3239.r.

```


Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0BAF0042D01_AF00293_1.
Location/Qualifiers

FEATURES

source
1. .843
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DB008YE02"
/tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 24.6%; Score 709.8; DB 13; Length 843;
Best Local Similarity 95.0%; Pred. No. 4.3e-109;
Matches 776; Conservative 0; Mismatches 34; Indels 7; Gaps 4;
1694 CAGCTGCACAGCGGAGAGCTGAGG--TTGCCCTAGTGTGCTGCAGACCTGCTG 1751
|||||
824 CACGCGCCAAAGCGGAGAGTGTAGGTTGCCCAACCGGGCTGTGACACCTGTTG 765
|||||
1752 AGTCCCATCTTAGTGTGACCCCTGGGCTGCTGCTCACCAGGCGAGTCATCTGGCT 1811
|||||
764 AGTCCCATCTTAGTGTGAA--CCCTGGGCTGCTGCTCACCAGGCGAGTCATCTGGT 706
|||||
1812 ATGACCACTGTGGGAGCGCCAGCCCTCAGAGCTGAGCCTGCGCCTC-AAAAAGCAGTC 1870
|||||
705 ATGACCACTGTGGGAGCGCCAGCCCTGACAGCTGAGCCTGCGCCTCAAAAAAGCAGTC 646
|||||
1871 GTGGAGGAGCTGGGAGCAGAGTGTGCTGACCTGGGCGAGAGGCGCCTCCACCT 1930
|||||
645 GTGGAGGAGCTGGG--AGTATGCTGCTGACCTGGGCGAGNAGCGCCTCCACCT 589
|||||
1931 CTACTGTCAGCTGAGGCGCAGTGTGCTGCTGCTGCTTACCGAGCAGCTGGGCCGCTT 1990
|||||
588 CTACTGTCAGCTGAGGCGCAGTGTGCTGCTGCTTACCAGAGCAGCTGGGCCGCTT 529
|||||
1991 TGCCCTGGTGGGAGGCGCCTCAGCTGGCTGCGCGCAAGCGCTCAAGCTGCTTCTGTT 2050
|||||
528 TGCCCTGGTGGGAGGCGCCTCAGCTGGCTGCGCGCAAGCGCTCAAGCTGCTTCTGTT 469
|||||
2051 TGCCCGGTCGCTGACCTCCCTCGAGTACAACTCGGGTCTACTGCTGCTGATGACAC 2110
|||||
468 TGCCCGGTCGCTGACCTCCCTCGAGTACAACTCGGGTCTACTGCTGCTGATGACAC 409
|||||
2111 CCAGATGCACTCAAGGAGTGGTGCAGCTGGAGAGCAGCTGGGGGAGCAGCTGATCCA 2170
|||||
408 CCAGATGCACTCAAGGAGTGGTGCAGCTGGAGAGCAGCTGGGGGAGCAGCTGATCCA 349
|||||
2171 GGAGCCACGGGCTCTGCACTTCAAGGACAGTTACCAAACTGCGCTTATCCATCCACGA 2230
|||||
348 GGAGCCACGGGCTCTGCACTTCAAGGACAGTTACCAAACTGCGCTTATCCATCCACGA 289
|||||
2231 TGTGCCAGCTCCCTGTTGGAGAGTACCTTGTGCTGCTTACAGAGATCCCTTTTA 2290
|||||
288 TGTGCCAGCTCCCTGTTGGAGAGTACCTTGTGCTGCTTACAGAGATCCCTTTTA 229
|||||
2291 TCACATCTGGAATGGCAGCGGCTACTTGTGCTGCTTACAGAGATCCCTTTTA 2350
|||||
228 TCACATCTGGAATGGCAGCGGCTACTTGTGCTGCTTACAGAGATCCCTTTTA 169
|||||
2351 CCCAGACCTAGTACCTGGCTGCAAGCTGTGGGTGTGGAGTGGAGGCGGCGGCA 2410
|||||
168 CCCAGACCTAGTACCTGGCTGCAAGCTGTGGGTGTGGAGTGGAGGCGGCGGCA 109
|||||
2411 GAGCTTCAGCATCACTTCAACATCACCAAGGACACAAAGGTTTCTGAGCTGCTGGCTCT 2470
|||||
108 GAGCTTCAGCATCACTTCAACATCACCAAGGACACAAAGGTTTCTGAGCTGCTGGCTCT 49
|||||

Qy 2471 GGAGAGTGAAGCGGGGTCCAGCCCTGCTGGGCCCC 2507
|||||
Db 48 GGAGAGTGAAGCGGGGTCCAGCCCTGCTGGGCCCC 12
|||||

RESULT 14

LOCUS BX345406
DEFINITION BX345406 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DC023YB10 5-PRIME, mRNA sequence.
ACCESSION BX345406
VERSION BX345406
KEYWORDS BX345406.1 GI:30369001
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

AUTHORS 1 (bases 1 to 939)
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3529.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0AC023DA05QPI&cluster=3529.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0AC023DA05QPI.

FEATURES

source
1. .939
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC023YB10"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 24.3%; Score 700.2; DB 13; Length 939;
Best Local Similarity 94.0%; Pred. No. 1.9e-107;
Matches 755; Conservative 22; Mismatches 20; Indels 6; Gaps 5;
Qy 1 AGCTGGGGCTCCGGGCTGAGCGCTAAAGCGGCTCCCGGCGGCGGCGGCGGCGGCGG 60
|||||
Db 142 AGCTGGGGCTCCGGGCTGAGCGCTAAAGCGGCTCCCGGCGGCGGCGGCGGCGGCGG 201
|||||
Qy 61 GCCCGCGGCTGCGCCCGCGGCGGCTGCGGCTGCGGCGGCGGCTGCGGCGGCGGCTG 120
|||||
Db 202 GCCCGCGGCTGCGCCCGCGGCGGCTGCGGCTGCGGCGGCGGCTGCGGCGGCGGCT 261
|||||
Qy 121 TGGGCATAGTCTGCGCGCTTGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 180
|||||
Db 262 TGGGCATAGTCTGCGCGCTTGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 321
|||||
Qy 181 CCACCCAGTGGCTGGTGCACAAACCGGACCTGCTCCCGCTTCTCTGCTGAGGCGGCGG 240
|||||
Db 322 CCACCCAGTGGCTGGTGCACAAACCGGACCTGCTTCTCTGCTGAGGCGGCGGCGG 381
|||||
Qy 241 ATGTGTACATCGTCAGACAAACAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
|||||
Db 382 ATGTGTACATGTTTAAAGACAAAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 439
|||||
Qy 301 AGATCTTCTCAAGTGCAACGGGGAGTGGGTGGCGGAGTGGGACCACTGATCGAGCGCA 360
|||||

digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

Db	440	AGATTTCTTAAAGTGC	AA-GTGGAGTGGGTGCGGCAGGTGACCAAGTATCGAGCGCA	498
Qy	361	GCACAGACGGGAGCAGTGGTGGAGCGGACCATGAGGTCCGCAATTAATGTCTCAAGGAGC	420	
Db	499	-SMCWACGGGAGCGTGGGTGCGCCACCATGGAGTCCGCAATTAATGTCTCAAGGAGC	557	
Qy	421	AGGTGAGAGAGTGTTCGGGCTGGAGGAACTACTGTGCCAGTGCCTGGATGAGTCTCT	480	
Db	558	AGGTGAGAGAGTGTTCGGGCTGGAGGAACTACTGTGCCAGTGCCTGGATGAGTCTCT	617	
Qy	481	CGGGACACACAGAGTCAAGAGGCTTACATCCGATAGCCAGATTGCGCAAGAACTTCG	540	
Db	618	CGGGACACACAGAGTCAAGAGGCTTACATCCGATAGCCATTCGCGAGAACTTCG	677	
Qy	541	AGCAGAGCGGTGGCCAAAGGAGTGTCCCTGGAGCAGGCGCATGCTGCTGCCCTGCCGTC	600	
Db	678	AGCAGAGCGGTGGCCAAAGGAGTGTCCCTGGAGCAGGCGCATGCTGCTGCCCTGCCGTC	737	
Qy	601	CACCGAGGAGTCCCTCCAGCGGAGGTGAGTGGCTCCGACAGGAGACCTGGTGGAC	660	
Db	738	CACCGAGGAGTCCCTCCAGCGGAGGTGAGTGGCTCCGACAGGAGACCTGGTGGAC	797	
Qy	661	CGTCCCTGGACCCCAATGATATACATCAGCGGAGCAGACGCTT-GGTGGTGGCAGAGGCC	719	
Db	798	CGTCCCTGGACCCCAATGATATACATCAGCGGAGCAGACGCTT-GGTGGTGGCAGAGGCC	857	
Qy	720	CGCTTGTCTGACACGGCCAACTACCTCGTGGTGGCCAAAGAACATCGTGGCACCTGCCGC	779	
Db	858	CGCTTGTCTGACACGGCCAACTACCTCGTGGTGGCCAAAGAACATCGTGGCACCTGCCGC	917	
Qy	780	AGCGCTTCCGCTGCTGTCATCGT	802	
Db	918	-GGGCTTCCGCTGCTGTCATCGT	939	
RESULT 15				
BX345407				
LOCUS				
DEFINITION				
BX345407 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens				
cDNA clone CS0DC023YB10 5-PRIME, mRNA sequence.				
BX345407				
BX345407.1 GI:30369002				
EST.				
Homo sapiens (human)				
ORGANISM				
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
1 (bases 1 to 1040)				
Li, W.B., Gruber, C., Jesse, J., and Polayes, D.				
Full-length cDNA libraries and normalization				
Unpublished (2001)				
Contact: Genoscope				
Genoscope - Centre National de Sequencage				
BP 191 91006 EVRY cedex - France				
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr				
Library was constructed by Life Technologies, a division of				
Invitrogen. This sequence belongs to sequence cluster 3529.f For				
more information about this cluster, see				
http://www.genoscope.cns.fr/				
cgi-bin/cluster.cgi?seq=CS0AC023DA05QP2&cluster=3529.f. Contact :				
Feng Liang Email : fliang@lifetech.com URL :				
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600				
Faraday Avenue Genoscope sequence ID : CS0AC023DA05QP2.				
Location/Qualifiers				
1..1040				
/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="taxon:9606"				
/clone="CS0DC023YB10"				
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"				
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"				
/note="1st strand cDNA was primed with a NotI-oligo (dT)				
primer. Five prime end enriched, double-strand cDNA was				

Search completed: October 6, 2004, 17:50:24
Job time : 7199 secs

THIS PAGE BLANK (USPTO)